

OM protein - protein search, using sw model

Run on: April 23, 2004, 14:55:02 ; Search time 72 Seconds
(without alignments)
6533.911 Million cell updates/sec

Title: US-10-697-263-2
Perfect score: 8740
Sequence: 1 MGCCRLGGCGSVSHSVSQ.....RNREKRALYKHNIAQVR 1665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1380s:*
2: Geneseqp1390s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1665	5 AAO15372	AAO15372 Human myo
2	8727	99.9	1665	5 AAE24151	AAE24151 Human kin
3	8423	96.4	2630	6 ABG76186	ABG76186 Human ser
4	8423	96.4	7968	6 ABG76187	ABG76187 Human ser
5	8417	96.3	2596	4 AAB30569	AAB30569 A splice
6	8403	96.1	1638	4 AAB30569	AAB30569 Human pro
7	8399	96.1	1620	4 AAB30568	AAB30568 A full le
8	7069	80.9	1351	4 AAB30567	AAB30567 Amino aci
9	4533	51.9	871	5 AAE16274	AAE16274 Human kin
10	4529	51.8	871	4 AAB30570	AAB30570 A splice
11	2447.5	28.0	548	4 AAB30571	AAB30571 A full le
12	2415.5	27.6	548	4 AAB30572	AAB30572 A full le
13	1973	22.6	2286	4 AAB65635	AAB65635 Novel pro
14	1973	22.6	2380	5 AAE19150	AAE19150 Human kin
15	1973	22.6	3208	7 ADE47676	ADE47676 Human NOV
16	1973	22.6	3267	7 ADC99075	ADC99075 Human KPP
17	1967.5	22.5	3268	7 ADB79959	ADB79959 Human kin
18	1931	22.1	3252	7 ADE47674	ADE47674 Human NOV
19	1913	21.9	2231	7 ADB79962	ADB79962 Human KIA
20	1864	21.3	3262	7 ADB79961	ADB79961 Mouse ser
21	1746.5	20.0	3186	7 ADE47672	ADE47672 Human NOV
22	632.5	7.2	282	7 ADE47678	ADE47678 Human NOV
23	607	6.9	127	5 AAB35501	AAB35501 Human kin
24	592.5	6.8	2861	2 AAW7227	AAW7227 Human TRI
25	592.5	6.8	2861	4 ABG17024	ABG17024 Novel hum

26	592.5	6.8	2861	4 AAG68192	AAG68192 GTPase pr
27	592.5	6.8	2861	6 ABR41116	ABR41116 Human GTP
28	592.5	6.8	2861	6 ABU70968	ABU70968 Human adi
29	592.5	6.8	2861	7 ADB98736	ADB98736 Human GTP
30	592.5	6.8	2861	7 AD82632	AD82632 Human pro
31	592.5	6.8	2861	7 AD899085	AD899085 TAR288.1
32	592.5	6.8	2861	7 AD899085	AD899085 Novel hum
33	592.5	6.8	2861	4 ABG06312	ABG06312 Novel hum
34	592.5	6.8	2861	4 ABG20756	ABG20756 Novel hum
35	592.5	6.8	3038	4 ABG17025	ABG17025 Novel hum
36	592.5	6.8	3038	6 ABM32684	ABM32684 Human Tri
37	592.5	6.8	3038	6 ABO07226	ABO07226 Human p53
38	592.5	6.8	3060	4 ABG06311	ABG06311 Novel hum
39	561.5	6.4	416	3 AAB53347	AAB53347 Human col
40	541	6.3	7107	4 ABB58144	ABB58144 Drosophil
41	536.5	6.1	1289	2 AAY27163	AAY27163 Peptide S
42	536.5	6.1	1289	3 AAY56781	AAY56781 Human Tra
43	529	6.1	913	4 ABB62810	ABB62810 Drosophil
44	529	6.1	913	4 ABB65968	ABB65968 Drosophil
45	489	5.6	448	2 AAY06922	AAY06922 Murine Zi

ALIGNMENTS

RESULT 1
AAO15372
ID AAO15372 standard; protein; 1665 AA.
XX
AC AAO15372;
XX
DT 19-SEP-2002 (first entry)
XX
DE Human myosin light chain kinase subfamily-related kinase protein.
XX
KW Human; gene therapy; chromosome 1; kinase protein;
KW myosin light chain kinase subfamily; kinase protein-mediated disease;
KW transgenic animal.
XX
OS Homo sapiens.
XX
PN WO2002040683-A2.
XX
PD 23-MAY-2002.
XX
PF 22-OCT-2001; 2001WO-US032616.
XX
PR 14-NOV-2000; 2000US-00711134.
PR 17-MAY-2001; 2001US-00858664.
(PEKE) PE CORP NY.
Wei M, Ketchum K, Di Francesco V, Beasley EM;
WFI; 2002-500223/53.
N-PSDB; AAL43908, AAL43909.
New kinase proteins related to myosin light chain kinase subfamily and encoding polynucleotide, useful for diagnosing, treating disease or condition mediated by the kinase protein and for identifying modulators.
Claim 1; Fig 2; 96pp; English.
The invention comprises the amino acid and coding sequences (located on chromosome 1) of a human kinase protein that is related to the myosin light chain kinase subfamily. The human kinase DNA and protein sequences of the invention are useful for identifying agents that modulate the activity of the human kinase protein. Kinase-modulating agents are useful for treating a disease or condition mediated by a human kinase protein. The human kinase DNA sequences can be used to produce transgenic animals which are useful for studying the function of kinase proteins and identifying/evaluating modulators of kinase protein activity. The present amino acid sequence represents the human kinase protein of the invention

Db	961	SSPEPTWEDIGQVSLVQIRIDLSGDAAADTISLDSIEVDPAYLNISDLVDIKYLFFEFM	1020
Qy	1021	IFRKVKSAQPEPPSPMAEEELAEFPPTWPMFEGELGPHAGLEITBESEVDALLAEAAV	1080
Db	1021	IFRKVKSAQPEPPSPMAEEELAEFPPTWPMFEGELGPHAGLEITBESEVDALLAEAAV	1080
Qy	1081	GRKKWSSPSSRLFPHRGRHLPLDEPAEILGRPVKASVEHISRIILKGRPEGLEKGGPPR	1140
Db	1081	GRKKWSSPSSRLFPHRGRHLPLDEPAEILGRPVKASVEHISRIILKGRPEGLEKGGPPR	1140
Qy	1141	KKPGCLSPRLSGLSKWDRAPTFLRELSDETVVLGQSVTLACQVSAQAPAAQATWSKDGAFL	1200
Db	1141	KKPGCLSPRLSGLSKWDRAPTFLRELSDETVVLGQSVTLACQVSAQAPAAQATWSKDGAFL	1200
Qy	1201	ESSSRVLISATLKNFOLLTIIVVAEDLGVYTCVSUNALGTVTTCGLVKAKERPSSPCP	1260
Db	1201	ESSSRVLISATLKNFOLLTIIVVAEDLGVYTCVSUNALGTVTTCGLVKAKERPSSPCP	1260
Qy	1261	DIEGVADGVLLVWKPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGTTYT	1320
Db	1261	DIEGVADGVLLVWKPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGTTYT	1320
Qy	1321	FRTACYSKAGMGYPYSPSEBOVLGGPSPHLASBEESQGRSAQPLPSTKTFATQIQGRGF	1380
Db	1321	FRTACYSKAGMGYPYSPSEBOVLGGPSPHLASBEESQGRSAQPLPSTKTFATQIQGRGF	1380
Qy	1381	SVTRQCKEASGALAAKIIPHPKDKTAVLREYEALKGLRPHILAQLHAAYLSPRHLVL	1440
Db	1381	SVTRQCKEASGALAAKIIPHPKDKTAVLREYEALKGLRPHILAQLHAAYLSPRHLVL	1440
Qy	1441	ILELCSGPPELLPCLAERASYSSEVKDYLWQMLSAQYLNHNOHILHLDLRSENMIITEYN	1500
Db	1441	ILELCSGPPELLPCLAERASYSSEVKDYLWQMLSAQYLNHNOHILHLDLRSENMIITEYN	1500
Qy	1501	LLKVVLDLGNQALSQSKVLPSDKFDYLETMAPELLEGOGAVPOTDIWAIGVTAFTMLSA	1560
Db	1501	LLKVVLDLGNQALSQSKVLPSDKFDYLETMAPELLEGOGAVPOTDIWAIGVTAFTMLSA	1560
Qy	1561	EYVPSSEGAARDLQRLGRLGLVRLSRCYAGLSGGAVAFIASTLTCAQPWGRPCASSCLQCFW	1620
Db	1561	EYVPSSEGAARDLQRLGRLGLVRLSRCYAGLSGGAVAFIASTLTCAQPWGRPCASSCLQCFW	1620
Qy	1621	LTEEGACSRPAPVPTPTARLRVFNVRKRRALLYKRHNLAQVR	1665
Db	1621	LTEEGACSRPAPVPTPTARLRVFNVRKRRALLYKRHNLAQVR	1665
RESULT 2			
AAE24151			
ID	AAE24151 standard; protein; 1665 AA.		
XX			
AC	AAE24151;		
XX			
DT	23-SEP-2002 (first entry)		
XX			
DE	Human kinase (PKIN)-22 protein.		
XX			
KW	Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;		
KW	acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;		
KW	asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;		
KW	development; hepatitis; cardiovascular; hypertension; drug screening;		
KW	myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;		
KW	fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;		
KW	hypercholesterolemia; obesity; gene therapy; cytostatic; anti-hiv;		
KW	neuroprotective; hepatotrophic; hypotensive; cardiant; nephrotropic;		
KW	hyperlipidaemia; enzyme.		
XX			
OS	Homo sapiens.		
XX			
FH	Key Location/Qualifiers		
FT	Domain 68..128		
FT	/note= "Immunoglobulin domain"		
FT	165..418		

FT Domain /note= "Eukaryotic protein kinase domain"
 FT 167..401
 FT /note= "Protein kinase domain"
 FT 1174..1235
 FT /note= "Immunoglobulin domain"
 FT 1369..1621
 FT /note= "Eukaryotic protein kinase domain"
 FT 1372..1606
 FT /note= "Protein kinase domain"
 XX WO200233099-A2.
 XX 25-APR-2002.
 XX 20-OCT-2001; 2001WO-US047728.
 XX 20-OCT-2000; 2000US-0242410P.
 XX 27-OCT-2000; 2000US-0244068P.
 XX 03-NOV-2000; 2000US-0245708P.
 XX 09-NOV-2000; 2000US-0247672P.
 XX 16-NOV-2000; 2000US-0249565P.
 XX 22-NOV-2000; 2000US-0252730P.
 XX 01-DEC-2000; 2000US-0250807P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Gururajan R, Baughn MR, Wadia NK, Elliott VS, Xu Y, Arvizu C;
 PI Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AdA, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;
 PI Recipon SA, Lu DAM, Borowsky ML, Thornton M, Swarnaker A;
 PI Thangavelu K, Khan FA, Ison CH;
 XX WPI: 2002-454603/48.
 XX N-PSDB; AAD38865.
 XX New human kinase polypeptide, for diagnosing, preventing and treating
 PT cancer, immune system disorders, growth and development disorders,
 PT cardiovascular disorders and lipid disorders.
 XX Claim 1: Page 182-186; 210pp; English.
 XX The invention relates human kinases (PKIN) and their corresponding
 CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
 CC treating and preventing cancer, an immune system disorder (e.g., acquired
 CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
 CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
 CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
 CC cardiovascular disorder (e.g., hypertension, myocardial infarction,
 CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
 CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,
 CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
 CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
 CC condition or a disease associated with the expression of PKIN in a
 CC biological sample. A composition comprising PKIN or an agonist or
 CC antagonist of PKIN is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional PKIN.
 CC PKIN is useful in a number of drug screening techniques and to analyse
 CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
 CC knockin humanised animals or transgenic animals to model human diseases,
 CC and in somatic or germline gene therapy. The present sequence is human
 CC PKIN protein
 XX SQ Sequence 1665 AA;
 Query Match 99.9%; Score 8727; DB 5; Length 1665;
 Best Local Similarity 99.9%; Pred No 0;
 Matches 1663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGCCRLGCGGCSVAHVSQGLTNHPSVMVCGWHFGLCGWGGGLHSSLPALPGPPSMQVTI 60
 DB 1 MGCCRLGCGGCSVAHVSQGLTNHPSVMVCGWHFGLCGWGGGLHSSLPALPGPPSMQVTI 60
 QY 61 EDVQAGTGTGTAQFEAIIIEGDPQPSVTWKSDVQLVDSRLSQOQEGTTSYLVLRHASKD 120

DB 61 EDVQAGTGTGTAQFEAIIIEGDPQPSVTWKSDVQLVDSRLSQOQEGTTSYLVLRHASKD 120
 QY 121 AGVYTCLAQNTGGQVLCKAELLVLGDNEPDSEKSHRKLHSHFVYKXEEIGRGVGFVK 180
 DB 121 AGVYTCLAQNTGGQVLCKAELLVLGDNEPDSEKSHRKLHSHFVYKXEEIGRGVGFVK 180
 QY 181 RVQHKGNKILCAAKFIPILRSRTRAQAYRERDILAAASHPLVTGLDQFETKTLILLEL 240
 DB 181 RVQHKGNKILCAAKFIPILRSRTRAQAYRERDILAAASHPLVTGLDQFETKTLILLEL 240
 QY 241 CSSELLDLRLYKGVVTEAEVKVYIQQVLEGLYHSHGVHLHDIKFSNIIIMVHPAREDI 300
 DB 241 CSSELLDLRLYKGVVTEAEVKVYIQQVLEGLYHSHGVHLHDIKFSNIIIMVHPAREDI 300
 QY 301 KICDFGFAQNTIPAELOFSQYSGSPFVSPEIIQONPVSEASDIWAMGVISYLSITCSPP 360
 DB 301 KICDFGFAQNTIPAELOFSQYSGSPFVSPEIIQONPVSEASDIWAMGVISYLSITCSPP 360
 QY 361 AGESDRATLLNVLEGRVSWSSPMAHLSADAKFIKATLQAPQAPSAACLSHPWFLK 420
 DB 361 AGESDRATLLNVLEGRVSWSSPMAHLSADAKFIKATLQAPQAPSAACLSHPWFLK 420
 QY 421 SMPAEAEHFINTKQLKFLARSWORSIMSYKSTILVMSIPEILRGPPDSPSLGVARHLK 480
 DB 421 SMPAEAEHFINTKQLKFLARSWORSIMSYKSTILVMSIPEILRGPPDSPSLGVARHLK 480
 QY 481 RTDGSSESSSSSSDNELAPPARAKSLPPSPVTHSPLHLHPRGFLRPPSASLPPEAEASRST 540
 DB 481 RTDGSSESSSSSSDNELAPPARAKSLPPSPVTHSPLHLHPRGFLRPPSASLPPEAEASRST 540
 QY 541 EAPAPPASPEGAGPPAAQGCVPFRHSVIRSLFYHQGESPEHGALAPGSRHRHARRHLK 600
 DB 541 EAPAPPASPEGAGPPAAQGCVPFRHSVIRSLFYHQGESPEHGALAPGSRHRHARRHLK 600
 QY 601 GGYIAGALPGLREPLMEHRVLEEAAREQATLLAKAPSPETALRLPASGTHLAPGSHS 660
 DB 601 GGYIAGALPGLREPLMEHRVLEEAAREQATLLAKAPSPETALRLPASGTHLAPGSHS 660
 QY 661 LEHDSPTSPRSSACGAQRLPSAPGAPIRDMGHPOGSKQLPSTGGHPTAQPRPS 720
 DB 661 LEHDSPTSPRSSACGAQRLPSAPGAPIRDMGHPOGSKQLPSTGGHPTAQPRPS 720
 QY 721 PDSPMGQAPFCHPKQSGAPQGGSPHAPVACPPGFPFGSCKEAPLVSPSSPPLGQQA 780
 DB 721 PDSPMGQAPFCHPKQSGAPQGGSPHAPVACPPGFPFGSCKEAPLVSPSSPPLGQQA 780
 QY 781 PPAPAKASPPLDKMGPGDLSLGRPKPGCSPPGSAQSSQVSSLRVSSQVGTPEG 840
 DB 781 PPAPAKASPPLDKMGPGDLSLGRPKPGCSPPGSAQSSQVSSLRVSSQVGTPEG 840
 QY 841 PSLDAEGHTQAEADLSSTPTLQRPQSVTWKPSLGGRGYAGVAGYGTAFGGDAGGM 900
 DB 841 PSLDAEGHTQAEADLSSTPTLQRPQSVTWKPSLGGRGYAGVAGYGTAFGGDAGGM 900
 QY 901 LQGGPMWARIAWVSQSEEEQEARAESSEQEARAESPLPOVSARVPVVGRAPTR 960
 DB 901 LQGGPMWARIAWVSQSEEEQEARAESSEQEARAESPLPOVSARVPVVGRAPTR 960
 QY 961 SSPEPTMEDIGQVSLVQIRDLSDAADAATISLIDISEVDPAYNLSDLYDIKYLPEEFM 1020
 DB 961 SSPEPTMEDIGQVSLVQIRDLSDAADAATISLIDISEVDPAYNLSDLYDIKYLPEEFM 1020
 QY 1021 IFRKVKSAQPEPPSPMAEEELAEFPETWPCGLGPHAGLEITTESEVDALLAEAAV 1080
 DB 1021 IFRKVKSAQPEPPSPMAEEELAEFPETWPCGLGPHAGLEITTESEVDALLAEAAV 1080
 QY 1081 GRKRWSSPSSLSLPHFPGRHLPDDEPAELGLRERVKASVEHISRLKGRPEGLEKGGPR 1140
 DB 1081 GRKRWSSPSSLSLPHFPGRHLPDDEPAELGLRERVKASVEHISRLKGRPEGLEKGGPR 1140
 QY 1141 KKPGLASFLSLGSLKSWDRAPTFLFELSDETVTLQSVTLACQVSAQPAADATWSKDGAPL 1200

Db 1141 KXGLASFLSLGKSWDRAPFTFLRELSDETIVVGLQSVTLACQVSAQPAQAQTSKDGAPL 1200
 QY 1201 ESSSRVLISATLKNFOLLTLVVAEDLGVVTCVSNALGVTTTGVLRKAERPSSPCP 1260
 Db 1201 ESSSRVLISATLKNFOLLTLVVAEDLGVVTCVSNALGVTTTGVLRKAERPSSPCP 1260
 QY 1261 DIGEYVADGVLLVWKPVSYSYGVVTVVQCSLEGGSWTTLASDIPDCCLYLSKLSRGTYT 1320
 Db 1261 DIGEYVADGVLLVWKPVSYSYGVVTVVQCSLEGGSWTTLASDIPDCCLYLSKLSRGTYT 1320
 QY 1321 FRTACVSKAGMGFPYSSPSEQLVLLGPGFSLASEEBSQGRSAQPLPSTTKTFAFQCIQGRF 1380
 Db 1321 FRTACVSKAGMGFPYSSPSEQLVLLGPGFSLASEEBSQGRSAQPLPSTTKTFAFQCIQGRF 1380
 QY 1381 SVVRCWKEKASGRALAANKIIPHPKDKTAVLREYALKGLRHPHQAQHAAYLSRPHVL 1440
 Db 1381 SVVRCWKEKASGRALAANKIIPHPKDKTAVLREYALKGLRHPHQAQHAAYLSRPHVL 1440
 QY 1441 LLELCGPELLPCLAEASYSSESEKQYLMQWLSATQYLHNQTHLHDLSENNIITEYN 1500
 Db 1441 LLELCGPELLPCLAEASYSSESEKQYLMQWLSATQYLHNQTHLHDLSENNIITEYN 1500
 QY 1501 LKVVLDLGNAGLSOEKVLPSDKKDYLETWAPLLEGGQAVPOTDIWAGTAFIMLSA 1560
 Db 1501 LKVVLDLGNAGLSOEKVLPSDKKDYLETWAPLLEGGQAVPOTDIWAGTAFIMLSA 1560
 QY 1561 EYPVSESGARDLQRLKGLVRLSRVAGLGGAVAFRLSTLCAQWGRPCASSCLQCPW 1620
 Db 1561 EYPVSESGARDLQRLKGLVRLSRVAGLGGAVAFRLSTLCAQWGRPCASSCLQCPW 1620
 QY 1621 LTEEGPACSRPAPVTFPTARLRVFRNREKERALLYKRNLAQVR 1665
 Db 1621 LTEEGPACSRPAPVTFPTARLRVFRNREKERALLYKRNLAQVR 1665

RESULT 3

ABG76186
 ID ABG76186 standard; protein; 2630 AA.

XX AC ABG76186;

XX DT 09-MAY-2003 (first entry)

XX DE Human serine/threonine or protein kinase 59079.

XX KW Human; enzyme; serine/threonine kinase; protein kinase; 59079;
 KW cardiovascular disease; heart failure; myocardial infarction;
 KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma; immunogen;
 KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
 KW haemolytic anaemia; cellular proliferative disorder; cancer;
 KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
 KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
 KW multiple sclerosis.

XX OS Homo sapiens.

XX PN US2002168742-A1.

XX PD 14-NOV-2002.

XX PF 15-FEB-2002; 2002US-00077130.

XX PR 15-FEB-2001; 2001US-0269201P.

XX RA (MILL-) MILLENNIUM PHARM INC.

XX PI Kapeller-Libermann R, Acton SL;

XX DR WPI; 2003-298729/29.

XX DR N-PSDB; ABX11641.

XX PT Novel isolated human protein kinase, designated 59079 or 12599

PT polypeptide, useful as diagnostic and therapeutic agents for preventing

PT cardiovascular diseases, proliferative disorders, and protein kinase disorders.
 PT XX Claim 8; Page 48-54; 119pp; English.
 PS CC The invention relates to an isolated human serine/threonine or protein
 CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
 CC comprising at least 85% identity to the nucleic acids appearing as
 CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
 CC of the kinases or their fragments. Also included are a non-human host
 CC cell containing the nucleic acids, an antibody specific for the proteins,
 CC identifying a compound which binds to the kinase (by contacting the
 CC kinase or a cell expressing the kinase with a test compound and
 CC determining whether the kinase binds to the test compound) and modulating
 CC the activity of kinase using the identified compound. The kinases and
 CC their encoding nucleic acids are useful as diagnostic and therapeutic
 CC agents for preventing a disease or condition associated with an aberrant
 CC or unwanted 59079 or 12599 activity in a subject, including
 CC cardiovascular diseases such as heart failure, and myocardial infarction;
 CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
 CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
 CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
 CC such as cancer; and protein kinase disorders such as autoimmune
 CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
 CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
 CC and disorders are included in the specification). The kinases, their
 CC encoding nucleic acids and antibodies are useful in screening assays,
 CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
 CC pharmacogenomics). The kinases and their encoding nucleic acids are
 CC useful as query sequences to perform a search against public databases to
 CC identify other family members or related sequences. The present sequence
 CC represents kinase 59079
 XX SQ Sequence 2630 AA;

Query Match 96.4%; Score 8423; DB 6; Length 2630;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 53 PPSQVTTIEDVQAGTGTAGTQFAEIIIEGDPQPSVTVTKDSVQLVDSTRLSQOEGTYSIV 112
 Db 1018 PPSQVTTIEDVQAGTGTAGTQFAEIIIEGDPQPSVTVTKDSVQLVDSTRLSQOEGTYSIV 1077
 QY 113 LRHVASKDAGVYTCIAQNTGGQVLCRAELLVGGDNEPDSEKSHRRKLSHFVEYKEEIG 172
 Db 1078 LRHVASKDAGVYTCIAQNTGGQVLCRAELLVGGDNEPDSEKSHRRKLSHFVEYKEEIG 1137
 QY 173 RGVGFGFVRVQHKGNKILCAAKFIPILRSRTAQAYRERDILAAALSHPLVTGLDQFETRK 232
 Db 1138 RGVGFGFVRVQHKGNKILCAAKFIPILRSRTAQAYRERDILAAALSHPLVTGLDQFETRK 1197
 QY 233 TLILILELCSSEELLDRLYRKGVVTEAEVKVYIQQLVEGLVHLHGVLRLDKPSNIML 292
 Db 1198 TLILILELCSSEELLDRLYRKGVVTEAEVKVYIQQLVEGLVHLHGVLRLDKPSNIML 1257
 QY 293 VHPAREDIKICDFGPAQNTTPAELOFSQYGSPEFVSPEIIQONPVSEASDIWANGVISYL 352
 Db 1258 VHPAREDIKICDFGPAQNTTPAELOFSQYGSPEFVSPEIIQONPVSEASDIWANGVISYL 1317
 QY 353 SLTCSPPFAGESDRATILNLVLEGRVSWSPMAHLSEDAKDFIKATLQAPQAPSAQC 412
 Db 1318 SLTCSPPFAGESDRATILNLVLEGRVSWSPMAHLSEDAKDFIKATLQAPQAPSAQC 1377
 QY 413 LSHPWFLKSNPABEAHPINTKQLKFLARSRWORSIMSYKSLVWRSIPELLRGPDSPS 472
 Db 1378 LSHPWFLKSNPABEAHPINTKQLKFLARSRWORSIMSYKSLVWRSIPELLRGPDSPS 1437
 QY 473 LGVARHLCRDTGGSS 532
 Db 1438 LGVARHLCRDTGGSS 1497
 QY 533 AEASERSTEFAPAPAPGEGPAAQGCYPRHSVLSLFYHQAGSPHGHGALAPGSRHP 592

1498 AEASERSTEAPAPASPEGAPPAQCGVPRHSVIRSLFVHQAGESPEHCALAPGSRHRP 1557
593 ARRHLLKGGYIAGALPGLREPLMEHRVLEBEAAREEQATILLAKAPSFETALRLPASGTH 652
1558 ARREHLLKGGYIAGALPGLREPLMEHRVLEBEAAREEQATILLAKAPSFETALRLPASGTH 1617
653 LAPGSHSLHDSSTPRPSSEACEAQRPLSPAGCAPIRDMGHPOGSKQLPSTGCHPG 712
1618 LAPGSHSLHDSSTPRPSSEACEAQRPLSPAGCAPIRDMGHPOGSKQLPSTGCHPG 1677
713 TAQPERSPSPWQOPAPFCHPKGSAPOBQCSHPHAPVAPCPGSPFPCKEAPLVPSS 772
1678 TAQPERSPSPWQOPAPFCHPKGSAPOBQCSHPHAPVAPCPGSPFPCKEAPLVPSS 1737
773 PFLGQOPAPPAPAKASPLDLSKMGPGDISLPRPKPGPCSPGASQASSQVSSLRVGS 832
1738 PFLGQOPAPPAPAKASPLDLSKMGPGDISLPRPKPGPCSPGASQASSQVSSLRVGS 1797
833 SQVGTPEGPSLDAEGWTQEAEDLSDSSTPTLQRPQEQVTMRKFSLGGRGVAGVAGYTGA 892
1798 SQVGTPEGPSLDAEGWTQEAEDLSDSSTPTLQRPQEQVTMRKFSLGGRGVAGVAGYTGA 1857
893 FGGDAGGMLGQPMWARTAMAVSSEEEQEAAREASQSEEQEAREASPLPQVSARPPV 952
1858 FGGDAGGMLGQPMWARTAMAVSSEEEQEAAREASQSEEQEAREASPLPQVSARPPV 1917
953 EVGRAPTRASSPEPTFWEDIGQVSLVQIRDLSDGDAEAADTISLDSVDPAYLNLSDLYDI 1012
1918 EVGRAPTRASSPEPTFWEDIGQVSLVQIRDLSDGDAEAADTISLDSVDPAYLNLSDLYDI 1977
1013 KYLPPEFMI FRKVPKSAQPPSPMAEELAEFPETWPGELGPHAGLEITESESDVD 1072
1978 KYLPPEFMI FRKVPKSAQPPSPMAEELAEFPETWPGELGPHAGLEITESESDVD 2037
1073 ALLAEAAVGRKKWSPPRSKSLFHPFGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 1132
2038 ALLAEAAVGRKKWSPPRSKSLFHPFGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 2097
1133 LEKEGPRKPKGLASFRISGLKSWDRAPFTFLRELSDEIVVLGOSVTLACQVSAQPAQAT 1192
2098 LEKEGPRKPKGLASFRISGLKSWDRAPFTFLRELSDEIVVLGOSVTLACQVSAQPAQAT 2157
1193 WSKOGAPLESSESRVLISATLKNFQLLTLLVVAEDLGVYTCVSNALGTVTITGVLRAE 1252
2158 WSKOGAPLESSESRVLISATLKNFQLLTLLVVAEDLGVYTCVSNALGTVTITGVLRAE 2217
1253 RPSSSPCPDIGEVYADGVLLVWKPVESYGPVTYIVQCSLEGSWTTLASDIFDCCLYTSK 1312
2218 RPSSSPCPDIGEVYADGVLLVWKPVESYGPVTYIVQCSLEGSWTTLASDIFDCCLYTSK 2277
1313 LSRGTYTFRCTVSKAGMPVSSSEQVLLGGPSHLASEESQGRSAQPLPSTKTFAQ 1372
2278 LSRGTYTFRCTVSKAGMPVSSSEQVLLGGPSHLASEESQGRSAQPLPSTKTFAQ 2337
1373 TQIQGRFSVVRQCEKASGRALAAKIIPYHPKDKTAVLREYALKGLRHPLHQAQHAAY 1432
2338 TQIQGRFSVVRQCEKASGRALAAKIIPYHPKDKTAVLREYALKGLRHPLHQAQHAAY 2397
1433 LSPRHLVLILELCSGPPELLPCLAEASVSESVKDYLMQMSATQYLNQHILHDLRSE 1492
2398 LSPRHLVLILELCSGPPELLPCLAEASVSESVKDYLMQMSATQYLNQHILHDLRSE 2457
1493 NMIIETYNLLKVDLGNASQSLQEKVLPSEKFKDYLETNAPELLEGQAGVPTDITWAGV 1552
2458 NMIIETYNLLKVDLGNASQSLQEKVLPSEKFKDYLETNAPELLEGQAGVPTDITWAGV 2517
1553 TAPIMLSAEPVSSSEGARDLQRLKGLVRLSRCTAGLSGGAVAFIRSTLCQAPGRPCA 1612
2518 TAPIMLSAEPVSSSEGARDLQRLKGLVRLSRCTAGLSGGAVAFIRSTLCQAPGRPCA 2577
1613 SSCLOCPWLTEGPACSRPAPVTFPTARLRVFNREKRALLYKRNLAQVR 1665

Db RESULT 4
ABG76187
ID ABG76187 standard; protein; 7968 AA.
XX
AC ABG76187;
XX
DT 09-MAY-2003 (first entry)
XX
DE Human serine/threonine or protein kinase 12599.
XX
KW Human; enzyme; serine/threonine kinase; protein kinase; 12599;
cardiovascular disease; heart failure; myocardial infarction;
KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma; immunogen;
KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
KW haemolytic anaemia; cellular proliferative disorder; cancer;
KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
KW peritiasis; inflammatory bowel disease; rheumatoid arthritis;
multiple sclerosis.
XX
OS Homo sapiens.
XX
PN US2002168742-A1.
XX
PD 14-NOV-2002.
XX
PF 15-FEB-2002; 2002US-00077130.
XX
PR 15-FEB-2001; 2001US-0269201P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Kapeller-Libermann R, Acton SL;
XX
DR WPI; 2003-298729/29.
XX
N-PSDB; ABX11642.
XX
PT Novel isolated human protein kinase, designated 59079 or 12599
polypeptide, useful as diagnostic and therapeutic agents for preventing
cardiovascular diseases, proliferative disorders, and protein kinase
disorders.
XX
PS Claim 8; Page 84-104; 119pp; English.
XX
CC The invention relates to an isolated human serine/threonine or protein
kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
comprising at least 85% identity to the nucleic acids appearing as
ABX11641 and ABX11642 or their complement, a naturally occurring variant
of the kinases or their fragments. Also included are a non-human host
cell containing the nucleic acids, an antibody specific for the proteins,
CC identifying a compound which binds to the kinase (by contacting the
kinase or a cell expressing the kinase with a test compound and
determining whether the kinase binds to the test compound) and modulating
the activity of kinase using the identified compound. The kinases and
their encoding nucleic acids are useful as diagnostic and therapeutic
agents for preventing a disease or condition associated with an aberrant
or unwanted 59079 or 12599 activity in a subject, including
cardiovascular diseases such as heart failure, and myocardial infarction;
disorders involving blood vessels such as atherosclerosis, and Kaposi's
sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
such as cancer; and protein kinase disorders such as autoimmune
disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
rheumatoid arthritis, and multiple sclerosis (many examples of diseases
and disorders are included in the specification). The kinases, their
encoding nucleic acids and antibodies are useful in screening assays,
CC detection assays (e.g. forensic biology), and predicting clinical trials and
diagnostic assays, prognostic assays, and monitoring clinical trials and
pharmacogenomics). The kinases and their encoding nucleic acids are
useful as query sequences to perform a search against public databases to
identify other family members or related sequences. The present sequence

PD 26-OCT-2000.
XX
PF 11-APR-2000; 2000MO-US009488.
XX
PR 16-APR-1999; 99US-0129553P.
XX
PA (SCIO-) SCIOS INC.
XX
XX Zeng W, Stanton L, Kong H;
PI
XX WPI; 2001-007013/01.
XX N-PSDB; AAC62287.
DR
XX Novel h19G5 polypeptides capable of regulating signal transduction and
PT exhibiting kinase activity useful for identifying antibodies to treat
PT cardiac diseases, and additional mediators of signal transduction.
XX
PS Claim 1; Page 68-74; 81pp; English.
XX
CC The present sequence represents a splice variant of human in signal
CC transduction polypeptide. The polypeptide is designated H19G5. The
CC protein is capable of regulating signal transduction and exhibits kinase
CC activity. The H19G5 transcript is expressed in the heart. H19G5
CC polypeptides and polynucleotides are useful for preventing or treating a
CC cardiac disease, such as congestive heart failure, dilated congestive
CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
CC mitral valve disease, aortic valve disease or tricuspid valve disease,
CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
CC arterial or renovascular hypertension, arteriosclerosis, atherosclerosis
CC and cardiac tumours in humans. The polypeptide is also useful for
CC detecting the expression of a protein capable of regulating signal
CC transduction or the expression of a protein capable of acting as a donor
CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
CC be used as probes for detecting discrete antigens expressed by tissue or
CC cell samples, and therefore used in humans for localization and
CC monitoring of microbial infection
XX
SQ Sequence 2596 AA;
Query Match 96.3%; Score 8417; DB 4; Length 2596;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1611; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 53 PPSNQVITDVQAGTGAFAEIIEDGPPQSVTWYKDSQVLDSTRLSQQEGTTSYL 112
Db PPSNQVITDVQAGTGAFAEIIEDGPPQSVTWYKDSQVLDSTRLSQQEGTTSYL 1043
QY 113 LRHVASKDAGVYTCLAQNTGGVLCXAEILLVGGDNEPDSKSHRRKLHSEYVKEIG 172
Db 1044 LRHVASKDAGVYTCLAQNTGGVLCXAEILLVGGDNEPDSKSHRRKLHSEYVKEIG 1103
QY 173 RGVEGFVYRQVHKGKILCAAKFIPURSTRQAQYRERDILAAALSHPLVTGLDQFETR 232
Db 1104 RGVEGFVYRQVHKGKILCAAKFIPURSTRQAQYRERDILAAALSHPLVTGLDQFETR 1163
QY 233 TLILILELCSSEBLLDRLYRKGWVTAEVKVIYQQLVEGLHVLHSHGVHLHDIKPSNITLM 292
Db 1164 TLILILELCSSEBLLDRLYRKGWVTAEVKVIYQQLVEGLHVLHSHGVHLHDIKPSNITLM 1223
QY 293 VHPAREDIKICDFGAQNTIPAELOFSQYGSPEFVSPEIIQONPVSEASDIWAMGVISYL 352
Db 1224 VHPAREDIKICDFGAQNTIPAELOFSQYGSPEFVSPEIIQONPVSEASDIWAMGVISYL 1283
QY 353 SLTCSPPFAGESDRATLLNVLGKSVSWSPMAHLSEDAKDFIKATLORAPQARSAAQC 412
Db 1284 SLTCSPPFAGESDRATLLNVLGKSVSWSPMAHLSEDAKDFIKATLORAPQARSAAQC 1343
QY 413 LSHPWFLKSNPABEAHFINTKQLKFLIARSQRWRSIMSYKSLVNRSTPEILLRGPSPS 472
Db 1344 LSHPWFLKSNPABEAHFINTKQLKFLIARSQRWRSIMSYKSLVNRSTPEILLRGPSPS 1403
QY 473 LGVARHLCRDTGGSSSSSSSSSNNELAPFARAKSLPPSVTHSPLLHPRGFLRPSASLPEE 532

Db 1404 LGVARHLCRDTGGSSSSSSSNNELAPFARAKSLPPSVTHSPLLHPRGFLRPSASLPEE 1463
QY 533 AEASERSTEAPPASPEGAPPAAGCQVPRHSVIRSLFYHQAGSPSPHAGALAPGSRHP 592
Db 1464 AEASERSTEAPPASPEGAPPAAGCQVPRHSVIRSLFYHQAGSPSPHAGALAPGSRHP 1523
QY 593 ARRHLLKGYIAGALPGLRLEPMHRLVEEBAAREOATLLAKAPSFETALRLPASGTH 652
Db 1524 ARRHLLKGYIAGALPGLRLEPMHRLVEEBAAREOATLLAKAPSFETALRLPASGTH 1583
QY 653 LAPGSHSHLEHDSPTPTSPSSACGEAQRLLPSAPSGGAPIRDMGHPOGSKQLPSTGGHPG 712
Db 1584 LAPGSHSHLEHDSPTPTSPSSACGEAQRLLPSAPSGGAPIRDMGHPOGSKQLPSTGGHPG 1643
QY 713 TAQPERPSDPSWGPAPAPCHPKQSGAPQEGCSPPHAPVAPCPGSPFPGSCKEAPLVSS 772
Db 1644 TAQPERPSDPSWGPAPAPCHPKQSGAPQEGCSPPHAPVAPCPGSPFPGSCKEAPLVSS 1703
QY 773 PFIQGOQAPAPAKASPPILDSRMGPGDISLPGRPKPCSSPGSASQSSQVSSLRVGS 832
Db 1704 PFIQGOQAPAPAKASPPILDSRMGPGDISLPGRPKPCSSPGSASQSSQVSSLRVGS 1763
QY 833 SOVGTEPGSLDAEGWTOAEOLDSSTTLQRPQPOVMTMRKPSLGGRGYAGVAGYGTFA 892
Db 1764 SOVGTEPGSLDAEGWTOAEOLDSSTTLQRPQPOVMTMRKPSLGGRGYAGVAGYGTFA 1823
QY 893 FGDAGGMLQGQPMWARIAMVQSSEEEQEARABSOSEQOEARABSPLPQVSARVP 952
Db 1824 FGDAGGMLQGQPMWARIAMVQSSEEEQEARABSOSEQOEARABSPLPQVSARVP 1883
QY 953 EVGRAPTRSPPTWEDIGQVSLQIRDLSDAEAAATISLIDISEVDPAYINISDLXYI 1012
Db 1884 EVGRAPTRSPPTWEDIGQVSLQIRDLSDAEAAATISLIDISEVDPAYINISDLXYI 1943
QY 1013 KYLPFEFMIFRKVPKSAQPEPPSPMAEELAEFFPPTWPFGEIAPHAGLITSESEVD 1072
Db 1944 KYLPFEFMIFRKVPKSAQPEPPSPMAEELAEFFPPTWPFGEIAPHAGLITSESEVD 2003
QY 1073 ALLAEAAVGRKRWKSPGRSLFHPFGRHLPLDEPAELGLRVRKASVEHISRLKGRPEG 1132
Db 2004 ALLAEAAVGRKRWKSPGRSLFHPFGRHLPLDEPAELGLRVRKASVEHISRLKGRPEG 2063
QY 1133 LKKEGPRKPKGLASFRISGLKSWDRAPTFLRELSDETIVLGOSVTLACQVSAQPAQAT 1192
Db 2064 LKKEGPRKPKGLASFRISGLKSWDRAPTFLRELSDETIVLGOSVTLACQVSAQPAQAT 2123
QY 1193 WSKDQAPLESSSRVLSATLKNFOLLTILVVVAEDLVYTCVSNALGTVTITGVLRKAE 1252
Db 2124 WSKDQAPLESSSRVLSATLKNFOLLTILVVVAEDLVYTCVSNALGTVTITGVLRKAE 2183
QY 1253 RPSSEPCPDIGBYADGVLLVWKPVSIGPVTYIVQCSLEGSSWTTLASDIFDCCLTSK 1312
Db 2184 RPSSEPCPDIGBYADGVLLVWKPVSIGPVTYIVQCSLEGSSWTTLASDIFDCCLTSK 2243
QY 1313 LSRGGTYTPTACVSKAGMPYSSPSEOVLLGGPSHLASEEESQORSAQPLPSTKTFAQ 1372
Db 2244 LSRGGTYTPTACVSKAGMPYSSPSEOVLLGGPSHLASEEESQORSAQPLPSTKTFAQ 2303
QY 1373 TQIQGRFVSVCQWKEKASGRALAAKIIPIYHPKOKTAVLREYEAALKGLRHPLHQAHAAY 1432
Db 2304 TQIQGRFVSVCQWKEKASGRALAAKIIPIYHPKOKTAVLREYEAALKGLRHPLHQAHAAY 2363
QY 1433 LSPRHLVLILELCSGPELLPCLAEASYSSEVDYLMQMLSATQYLHNOHILHLDRSE 1492
Db 2364 LSPRHLVLILELCSGPELLPCLAEASYSSEVDYLMQMLSATQYLHNOHILHLDRSE 2423
QY 1493 NMIIETEYNLLKVVLDIGNAQSLSOEKVLPDSKFKDYLETMAPELLEGGQAVPQTDIWAIGV 1552
Db 2424 NMIIETEYNLLKVVLDIGNAQSLSOEKVLPDSKFKDYLETMAPELLEGGQAVPQTDIWAIGV 2483
QY 1553 TAFIMLSAEYPVSSSEGCARDLQRLKGLVRLSRVCYAGLSGGAVAFRLSTLCAQPMWRPCA 1612
Db 2484 TAFIMLSAEYPVSSSEGCARDLQRLKGLVRLSRVCYAGLSGGAVAFRLSTLCAQPMWRPCA 2543

QY 1613 SSCLCQPLWTEBGPACSRPAPVTFPTARLRFVVRNREKRALLYKHNLAQVR 1665
DB 2544 SSCLCQPLWTEBGPACSRPAPVTFPTARLRFVVRNREKRALLYKHNLAQVR 2596

RESULT 6
AAB85504
ID AAB85504 standard; protein; 1618 AA.
AC AAB85504;
XX 25-SEP-2001 (first entry)
XX Human protein kinase SGK145.
XX
KW Protein kinase; enzyme; cytostatic; nootropic; neuroprotective; human;
KW antiparkinsonian; virucide; antibacterial; antifungal; antimigraine;
KW analgesic; hypotensive; hypertensive; immunosuppressive; antiallergic;
KW antipsoriatic; antirheumatic; antiarthritic; ophthalmologic; anorectic;
KW osteopathic; thrombolytic; antiarteriosclerotic; antiasthmatic;
KW vasotropic; antidiabetic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155356-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US002337.
XX
PR 25-JAN-2000; 2000US-0178078P.
PR 31-JAN-2000; 2000US-0179364P.
PR 17-FEB-2000; 2000US-0183173P.
PR 17-MAR-2000; 2000US-0190162P.
PR 29-MAR-2000; 2000US-0193404P.
PR 13-NOV-2000; 2000US-0247013P.
XX
PA (SUGS-) SUGEN INC.
XX
PI Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX
XX WPI; 2001-476202/51.
DR N-PSDB; AAH45904.
XX
XX Kinase polypeptides useful for treating cancers, Alzheimer's disease,
PT viral infections, diabetes, obesity, organ transplant rejection and
PT rheumatoid arthritis.
XX
PS Claim 7; Page 215; 218pp; English.
XX
XX The invention provides human protein kinases and protein kinase-like
CC enzymes and polynucleotides encoding the polypeptides. The kinase
CC polypeptides and their modulators are useful for treating a disease or
CC disorder such as cancer, immune-related diseases, cardiovascular disease,
CC brain or neuronal-associated disease and metabolic disorders, including
CC cancers of tissues, cancers of hematopoietic origin, diseases of the
CC central nervous system, diseases of the peripheral nervous system,
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
CC lateral sclerosis, viral infections, infections caused by prions,
CC bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,
CC mood disorders, attention disorders, cognition disorders, hypotension,
CC hyperextension, psychotic disorders, neurological disorders, dyskinesias,
CC metabolic disorders, and organ transplant rejection. They are also useful
CC for treating rhinitis, autoimmunity, atherosclerosis, psoriasis,
CC osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic
CC inflammatory bowel disease, rheumatoid arthritis, metabolic disorders
CC such as diabetes, obesity, cardiovascular diseases such as reperfusion
CC injury, coronary thrombosis, clotting disorders and atherosclerosis,
CC ocular diseases such as glaucoma, retinopathy and macular degeneration,
CC psychiatric and neurological disorders such as anxiety, schizophrenia,
CC dementia, manic depression, etc. The polynucleotides are useful in gene
CC therapy techniques to treat the above mentioned disorders. Sequences

CC	AAB85491-85522	represent the human protein kinases of the invention
XX	Sequence 1618 AA;	
XX	Query Match	96.1%; Score 8403; DB 4; Length 1618;
XX	Best Local Similarity	99.6%; Pred. No. 0;
XX	Matches 1611; Conservative	0; Mismatches 1; Indels 6; Gaps 1;
QY	54	PSMQVTIEDVQAQCTGTAQFAEIIIEGDPQPSVTVYKDSVQLVDSVTRLSQQQGGTTTSLVL 113
DB	1	PSMQVTIEDVQAQCTGTAQFAEIIIEGDPQPSVTVYKDSVQLVDSVTRLSQQQGGTTTSLVL 60
QY	114	RHVASKDAGVVTCTLAQNTGGQVLCCKAELLVLG-----GDNEPDSEKSHRRKLHSFYEV 167
DB	61	RHVASKDAGVVTCTLAQNTGGQVLCCKAELLVLGAAASHSLGDNEDPSEKSHRRKLHSFYEV 120
QY	168	KEEIGRGVGFVKVVOHKGKILCAAKFIPLRSETRACAYEEDIIAALSHPLVTGLDQ 227
DB	121	KEEIGRGVGFVKVVOHKGKILCAAKFIPLRSETRACAYEEDIIAALSHPLVTGLDQ 180
QY	228	FETRTLLILIELSCSEBELLDRLYRKGVVTEAEVKVYIQQLVEGLHYLHSHGVLDIKP 287
DB	181	FETRTLLILIELSCSEBELLDRLYRKGVVTEAEVKVYIQQLVEGLHYLHSHGVLDIKP 240
QY	288	SNILMVHPAREDIKICDFGFAQNTTPAELOFSQYGSPEFVSPEIIQONPVSEASDIWAMG 347
DB	241	SNILMVHPAREDIKICDFGFAQNTTPAELOFSQYGSPEFVSPEIIQONPVSEASDIWAMG 300
QY	348	VISYLSLTCSPPFAGESDRATLLNVLEGRVSWSPMAAHLSEDAKOFIKATLQAPQARP 407
DB	301	VISYLSLTCSPPFAGESDRATLLNVLEGRVSWSPMAAHLSEDAKOFIKATLQAPQARP 360
QY	408	SAAQCLSHPWFLKSNPAEAHFIINTKQLKFLARSWORSLMSYKSLVWRS:PELLRGP 467
DB	361	SAAQCLSHPWFLKSNPAEAHFIINTKQLKFLARSWORSLMSYKSLVWRS:PELLRGP 420
QY	468	PDPSLGLVARHLCDRTGSGSS 527
DB	421	PDPSLGLVARHLCDRTGSGSS 480
QY	528	SUPEAEASERSTEAPAPSPAGAGPAAQGCVPRAHSVIRSLFYHQAGSPESHGALAPG 587
DB	481	SUPEAEASERSTEAPAPSPAGAGPAAQGCVPRAHSVIRSLFYHQAGSPESHGALAPG 540
QY	588	SRHPARRRHLLKGGYTAGALPGLREPLMEHRVLEBEAAEEQATLLAKAPSETALRLP 647
DB	541	SRHPARRRHLLKGGYTAGALPGLREPLMEHRVLEBEAAEEQATLLAKAPSETALRLP 600
QY	648	ASGTHLAPGHSHSLHSDSPSTPRPSSSEACGEAQLPSAPSGGAPIRDMGHPOGSKOLPST 707
DB	601	ASGTHLAPGHSHSLHSDSPSTPRPSSSEACGEAQLPSAPSGGAPIRDMGHPOGSKOLPST 560
QY	708	GGHGTATQEPSPDSKMGQAPAPCHPKQGSAPQEGCSHPAVAPCPGSPFGSCKEAP 767
DB	661	GGHGTATQEPSPDSKMGQAPAPCHPKQGSAPQEGCSHPAVAPCPGSPFGSCKEAP 720
QY	768	LVPSSPFLGQAPAPAPAKASPPLDKMGFGDISLFCRKPFGCSPFGSASQASSQVSS 827
DB	721	LVPSSPFLGQAPAPAPAKASPPLDKMGFGDISLFCRKPFGCSPFGSASQASSQVSS 780
QY	828	LRVGSQVGTGPGSLDAEGWTEADLSDTTLQRPQOVTMRKFSLGGRGGYAGVAG 887
DB	781	LRVGSQVGTGPGSLDAEGWTEADLSDTTLQRPQOVTMRKFSLGGRGGYAGVAG 840
QY	888	YGTAFGADAGMVGQGMWARIAMAVSQSEEEQEARAEASQSEEQQBARAESPLPQVS 947
DB	841	YGTAFGADAGMVGQGMWARIAMAVSQSEEEQEARAEASQSEEQQBARAESPLPQVS 900
QY	948	ARPYVEVGRAPTRSSPEPTWEDIGVSLVQIRDLSDGDAEADTISLDISEVDPAVLNL 1007
DB	901	ARPYVEVGRAPTRSSPEPTWEDIGVSLVQIRDLSDGDAEADTISLDISEVDPAVLNL 960
QY	1008	DLVDIKYLPFFMTFRKVPKSAQPEPPSPMAESELAEFPPTWPNPCELPHAGLEITEE 1067

Db 961 DLYDIKYLPEEFMIFKVPKSAQPPSPMAEEAEFPETWPGELQPHAGLSITEE 1020
 Qy 1068 SEDVDALLAAVGRKRWSSPSRSLFHPGPHLPDPAELGLRERVKASVSHISILK 1127
 Db 1021 SEDVDALLAAVGRKRWSSPSRSLFHPGPHLPDPAELGLRERVKASVSHISILK 1080
 Qy 1128 GRPEGLEKEGPPKPKPLGLASFLSLGLKSWDRAPTELRSLDETTLVGLGSVTLCAQVSAOP 1187
 Db 1081 GRPEGLEKEGPPKPKPLGLASFLSLGLKSWDRAPTELRSLDETTLVGLGSVTLCAQVSAOP 1140
 Qy 1188 AQAQWSDKCAPLESSRVLISATLKNFOLLITLWVAEDLGVTTCVSVALGTVTITGV 1247
 Db 1141 AQAQWSDKCAPLESSRVLISATLKNFOLLITLWVAEDLGVTTCVSVALGTVTITGV 1200
 Qy 1248 LRKAERPSSPCPDIGEVYADGVLLVWKPVESYGVTVIVQCSLEGGSWTTLASDIFDC 1307
 Db 1201 LRKAERPSSPCPDIGEVYADGVLLVWKPVESYGVTVIVQCSLEGGSWTTLASDIFDC 1260
 Qy 1308 YLTSKLSRGTYTFRACVSKAGMPYSPSEOVLLGSPHLSAEESSQGRSAQPLPSTK 1367
 Db 1261 YLTSKLSRGTYTFRACVSKAGMPYSPSEOVLLGSPHLSAEESSQGRSAQPLPSTK 1320
 Qy 1368 TFAFQTIQGRFVSVRVCWEKASGRALAAKIIIPYHPKDTAVLRREYALKGLRHPLAQ 1427
 Db 1321 TFAFQTIQGRFVSVRVCWEKASGRALAAKIIIPYHPKDTAVLRREYALKGLRHPLAQ 1380
 Qy 1428 LHAAYLSRHLVLLLELCSGPPELLCLAEARASYSESEVKDYLWOMLSATQYLHNOHILH 1487
 Db 1381 LHAAYLSRHLVLLLELCSGPPELLCLAEARASYSESEVKDYLWOMLSATQYLHNOHILH 1440
 Qy 1488 DLRSNNMITEYNLLKVDLGNALSOEKVLPDSKFKDYLETWAPLEGGQAVPQTDI 1547
 Db 1441 DLRSNNMITEYNLLKVDLGNALSOEKVLPDSKFKDYLETWAPLEGGQAVPQTDI 1500
 Qy 1548 WAIGVTAFIMLSAEPVSESGARDLQRLKGLVRLSRVAGLSGGAVAFRLSTLCAQPW 1607
 Db 1501 WAIGVTAFIMLSAEPVSESGARDLQRLKGLVRLSRVAGLSGGAVAFRLSTLCAQPW 1560
 Qy 1608 GRPCASSCLOCPWLEEGPACSRPAPVTFPTARLVFVRNREKRALLYKRNHIAQVR 1665
 Db 1561 GRPCASSCLOCPWLEEGPACSRPAPVTFPTARLVFVRNREKRALLYKRNHIAQVR 1618
 RESULT 7
 AAB30568
 ID AAB30568 standard; protein; 1610 AA.
 AC AAB30568;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE A full length human signal transduction polypeptide.
 XX
 KW Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
 KW congestive heart failure; dilated congestive cardiomyopathy;
 KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
 KW mitral valve disease; aortic valve disease; tricuspid valve disease;
 KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
 KW atherosclerosis; cardiac tumour; microbial infection.
 XX
 OS Homo sapiens.
 XX
 PN W0200063381-A1.
 XX
 PD 26-OCT-2000.
 XX
 PF 11-APR-2000; 2000WO-US009488.
 XX
 PR 16-APR-1999; 99US-0129553P.
 XX
 PA (SCIO-) SCIOS INC.
 XX

PI Zeng W, Stanton L, Kong H;
 XX WPI; 2001-007013/01.
 DR N-PSDB; AAC62286.
 XX
 PT Novel h19G5 polypeptides capable of regulating signal transduction and
 PT exhibiting kinase activity useful for identifying antibodies to treat
 PT cardiac diseases, and additional mediators of signal transduction.
 XX
 XX Claim 1; Page 61-65; 81pp; English.
 XX
 CC The present sequence represents a human protein with putative function in
 CC signal transduction. The polypeptide is designated H19G5. The protein is
 CC capable of regulating signal transduction and exhibits kinase activity.
 CC The H19G5 transcript is expressed in the heart. H19G5 polypeptides and
 CC polynucleotides are useful for preventing or treating a cardiac disease,
 CC such as congestive heart failure, dilated congestive cardiomyopathy,
 CC hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve
 CC disease, aortic valve disease or tricuspid valve disease, angina
 CC pectoris, myocardial infarction, cardiac arrhythmia, pulmonary, arterial
 CC or renovascular hypertension, arteriosclerosis, atherosclerosis and
 CC cardiac tumours in humans. The polypeptide is also useful for detecting or
 CC the expression of a protein capable of regulating signal transduction or
 CC molecule of a phosphate group. The monoclonal antibodies can be used as
 CC probes for detecting discrete antigens expressed by tissue or cell
 CC samples, and therefore used in humans for localization and monitoring of
 CC microbial infection
 XX
 XX Sequence 1610 AA;
 SQ
 Query Match 96.1%; Score 8399; DB 4; Length 1610;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1608; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 56 MOVTTEDVQATGGAQFAEIIIEGDPOPSVTWYKDSVQLVDSTRLSQOQEGTYSVLVLRH 115
 Db 1 MOVTTEDVQATGGAQFAEIIIEGDPOPSVTWYKDSVQLVDSTRLSQOQEGTYSVLVLRH 60
 Qy 116 VASKAGVVTCLAQNTGGQVLCBAELLVLDGNDPESEKSHRRKLMSPFVKEETGRGV 175
 Db 61 VASKAGVVTCLAQNTGGQVLCBAELLVLDGNDPESEKSHRRKLMSPFVKEETGRGV 120
 Qy 176 FGFVKRVQHKGNKILCAAKFIPLSRTRAQAYRERDILAAALSHPLVTGLDQFETKTLI 235
 Db 121 FGFVKRVQHKGNKILCAAKFIPLSRTRAQAYRERDILAAALSHPLVTGLDQFETKTLI 180
 Qy 236 LILELCSSEELLDRLYRKGVTVEAKVYIQQVLEGLVHLHSHGVLEHLDIKPSNIMLVHP 295
 Db 181 LILELCSSEELLDRLYRKGVTVEAKVYIQQVLEGLVHLHSHGVLEHLDIKPSNIMLVHP 240
 Qy 296 AREDIKICDFQAQNTIPAELOFSQYGSPEFVSPPIIQQNPVSEASDIWAMGVISYLSLT 355
 Db 241 AREDIKICDFQAQNTIPAELOFSQYGSPEFVSPPIIQQNPVSEASDIWAMGVISYLSLT 300
 Qy 356 CSPPAGESDRATLLNVLEGRVSWSSPMAHLSEDAKDFIKATLQAPQAPSAACLSH 415
 Db 301 CSPPAGESDRATLLNVLEGRVSWSSPMAHLSEDAKDFIKATLQAPQAPSAACLSH 360
 Qy 416 PWFLLKMPABEAHFINTKQLKFLARSRWORSLSVYKSLVWRSIPELLRGPPDPSLGV 475
 Db 361 PWFLLKMPABEAHFINTKQLKFLARSRWORSLSVYKSLVWRSIPELLRGPPDPSLGV 420
 Qy 476 ARHLCRDTGGSS 535
 Db 421 ARHLCRDTGGSS 480
 Qy 536 SERSTEAPAPASPEGAGPPAAQCVPRHSVIESLPHONGESPHEGALAGSRHRPARR 595
 Db 481 SERSTEAPAPASPEGAGPPAAQCVPRHSVIESLPHONGESPHEGALAGSRHRPARR 540
 Qy 596 RHLLKGVYIAGALPGLREPLMEHVRLEEEAAREEQATLLAKAPSPETALURLPASGTHLAP 655

541 RHLLKGGYIAGALFGLRPLMEHVRLEEAAREEQATLLAKAPSFETALRLPASGTHLAP 600
 QY 656 GHSLSLEHDSSTPRPSEACGEAQRLLPSAPSGGAPIRDMGHQPQSKQLPSTGGHPGTAQ 715
 Db 601 GHSLSLEHDSSTPRPSEACGEAQRLLPSAPSGGAPIRDMGHQPQSKQLPSTGGHPGTAQ 660
 QY 716 PERSPDQWOPAPFCHPKQGSAPQGCSPHFAVAPCPGSPFPPQSCKEAPLVSSPFL 775
 Db 661 PERSPDQWOPAPFCHPKQGSAPQGCSPHFAVAPCPGSPFPPQSCKEAPLVSSPFL 720
 QY 776 GQOQAPPAKASPPPLDSKMGPGDISLPGKPKGCPSSPGSASQASSQVSSLRVGSQV 835
 Db 721 GQOQAPPAKASPPPLDSKMGPGDISLPGKPKGCPSSPGSASQASSQVSSLRVGSQV 780
 QY 836 GTEPGPSLDAGTQOEADLSDSTPLQRPQEQVTVWRKSLGEGGYAGVAGYGFARGG 895
 Db 781 GTEPGPSLDAGTQOEADLSDSTPLQRPQEQVTVWRKSLGEGGYAGVAGYGFARGG 840
 QY 896 DAGMGLGQPMWARIAMAVSSEEEQEEARAEASQSEEEQEEARAEPLQVSGARVPEVG 955
 Db 841 DAGMGLGQPMWARIAMAVSSEEEQEEARAEASQSEEEQEEARAEPLQVSGARVPEVG 900
 QY 956 RAPTRSSPEPTWEDIGQVSLVQIRDSGDAEADTISLDSRVDPAVNLSDLDYDKYL 1015
 Db 901 RAPTRSSPEPTWEDIGQVSLVQIRDSGDAEADTISLDSRVDPAVNLSDLDYDKYL 960
 QY 1016 PFEMIFKVKPSAQPPPPPMABEELAEPEPTWMPGELPHAGLEITESEEDVDALL 1075
 Db 961 PFEMIFKVKPSAQPPPPPMABEELAEPEPTWMPGELPHAGLEITESEEDVDALL 1020
 QY 1076 AEAAGRRKWKSSPSRSLFHPGGRHLPLDEPAELGLRERVKASVEHISILKGRPEGLEK 1135
 Db 1021 AEAAGRRKWKSSPSRSLFHPGGRHLPLDEPAELGLRERVKASVEHISILKGRPEGLEK 1080
 QY 1136 EGPPKPKGLASFLSLGLSKWDRAPTFLRELSDETIVVVGOSVTLACQVSAQAQAATWSK 1195
 Db 1081 EGPPKPKGLASFLSLGLSKWDRAPTFLRELSDETIVVVGOSVTLACQVSAQAQAATWSK 1140
 QY 1196 DGAPLESSRVLISATLKNFOLLITLVVAEDLGVTCTSVNALGVTITGVLRKAERPS 1255
 Db 1141 DGAPLESSRVLISATLKNFOLLITLVVAEDLGVTCTSVNALGVTITGVLRKAERPS 1200
 QY 1256 SSPCDIGENVADGVLVWKPVEGYGVTVIVQCSLEGSSWTTLASDIFDCCYLTKLSR 1315
 Db 1201 SSPCDIGENVADGVLVWKPVEGYGVTVIVQCSLEGSSWTTLASDIFDCCYLTKLSR 1260
 QY 1316 GGTFTFRACVSKAGMPYSPSPSQVLLGGPSHLASEESQGRSAQPLPSTTKTFAQTQI 1375
 Db 1261 GGTFTFRACVSKAGMPYSPSPSQVLLGGPSHLASEESQGRSAQPLPSTTKTFAQTQI 1320
 QY 1376 QRGRESVYQWKEKASGALAAKIIYPHPKDKTAVILREYALKGLRPHLAOLHAAYLSP 1435
 Db 1321 QRGRESVYQWKEKASGALAAKIIYPHPKDKTAVILREYALKGLRPHLAOLHAAYLSP 1380
 QY 1436 RHLVLILELCSGPPELLPCLAEASYSSEVKYLMQMLSATQYLHNQHLHLDLSENMI 1495
 Db 1381 RHLVLILELCSGPPELLPCLAEASYSSEVKYLMQMLSATQYLHNQHLHLDLSENMI 1440
 QY 1496 ITEYNLLKVDLGNQAQSLQSKVLPSDKFDYLETWAPELLEGGGAVPQTDIWAIGVTAF 1555
 Db 1441 ITEYNLLKVDLGNQAQSLQSKVLPSDKFDYLETWAPELLEGGGAVPQTDIWAIGVTAF 1500
 QY 1556 IMLSLEYVSSSEGARDLQRLKGLVRLSRCYAGLSGGAFAFLRSTLCQPMGRPCASSC 1615
 Db 1501 IMLSLEYVSSSEGARDLQRLKGLVRLSRCYAGLSGGAFAFLRSTLCQPMGRPCASSC 1560
 QY 1616 LQCPWLTEGPACSRPAPVPTPTARLVFVVRNEKRRALLYKHNLAQVR 1665
 Db 1561 LQCPWLTEGPACSRPAPVPTPTARLVFVVRNEKRRALLYKHNLAQVR 1610

RESULT 8
 AAB30567

AAB30567 standard; protein; 1351 AA.
 AAB30567;
 19-MAR-2001 (first entry)
 Amino acid sequence of a human signal transduction polypeptide.
 Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
 congestive heart failure; dilated congestive cardiomyopathy;
 hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
 mitral valve disease; aortic valve disease; tricuspid valve disease;
 myocardial infarction; cardiac arrhythmia; arteriosclerosis;
 atherosclerosis; cardiac tumour; microbial infection.
 Homo sapiens.
 WO200063381-A1.
 26-OCT-2000.
 11-APR-2000; 2000WO-US009488.
 16-APR-1999; 99US-0129553P.
 (SCIO-) SCIOS INC.
 Zeng W, Stanton L, Kong H;
 WPI; 2001-007013/01.
 N-PSDB; AAC62285.
 Novel h19G5 polypeptides capable of regulating signal transduction and
 exhibiting kinase activity useful for identifying antibodies to treat
 cardiac diseases, and additional mediators of signal transduction.
 Claim 1; Page 55-57; 81pp; English.
 The present sequence represents a human protein with putative function in
 signal transduction. The polypeptide is designated H19G5. The protein is
 capable of regulating signal transduction and exhibits kinase activity.
 The H19G5 transcript is expressed in the heart. H19G5 polypeptides and
 polynucleotides are useful for preventing or treating a cardiac disease,
 such as congestive heart failure, dilated congestive cardiomyopathy,
 hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve
 disease, aortic valve disease or tricuspid valve disease, angina
 pectoris, myocardial infarction, cardiac arrhythmia, pulmonary, arterial
 or renovascular hypertension, arteriosclerosis, atherosclerosis and
 cardiac tumours in humans. The polypeptide is also useful for detecting
 the expression of a protein capable of regulating signal transduction or
 the expression of a phosphate group. The monoclonal antibodies can be used as
 probes for detecting discrete antigens expressed by tissue or cell
 samples, and therefore used in humans for localization and monitoring of
 microbial infection
 Sequence 1351 AA;
 Query Match 80.9%; Score 7069; DB 4; Length 1351;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1348; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 316 LPSQVGSPEFVSPRIIQONPVSEASDIWANGVSYLSLTCSSPFAGESDRAILLNVLEG 375
 Db 2 VQFSQVGSPEFVSPRIIQONPVSEASDIWANGVSYLSLTCSSPFAGESDRAILLNVLEG 61
 QY 376 RYSWSSPMAAHLSEDAKDFIKATLQAPQAPSAACQLSHPWFLKSPAEAEHFINTKQL 435
 Db 62 RYSWSSPMAAHLSEDAKDFIKATLQAPQAPSAACQLSHPWFLKSPAEAEHFINTKQL 121
 QY 436 KFLARSRWQRSLMSYKILVWRSIPPELLRPPDPSLGVARHLCRTGSSSSSSSDN 495
 Db 122 KFLARSRWQRSLMSYKILVWRSIPPELLRPPDPSLGVARHLCRTGSSSSSSSDN 181

1576 LRKGLVRLSRCYAGLSGGAVAFRLSTLCAOPWGRPCASSCLQCQWLTEEGPACSRPAPVT 1635
 1262 LRKGLVRLSRCYAGLSGGAVAFRLSTLCAOPWGRPCASSCLQCQWLTEEGPACSRPAPVT 1321
 1636 FFTAALRVFVRNREKRALLKRNLAQVR 1665
 1322 FFTAALRVFVRNREKRALLKRNLAQVR 1351

RESULT 9
 AAE16274
 ID AAE16274 standard; protein; 871 AA.
 XX AAE16274;
 XX AC
 XX DT 26-MAR-2002 (first entry)
 XX DE Human kinase PKIN-20 protein.
 XX KW Human; kinase; PKIN-20; cancer; leukaemia; adenocarcinoma; osteoporosis;
 KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
 KW Acquired immune deficiency syndrome; AIDS; Addison's disease; anaemia;
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
 KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
 KW congestive heart failure; ischaemic heart disease; lung tumour; gout;
 KW fatty liver; Niemann-Pick's disease; gene therapy.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Domain 575..827
 FT /note="Eukaryotic protein kinase domain"
 FT Domain 580..812
 FT /label="Protein_kinase_domain"
 XX WO200196547-A2.
 XX 20-DEC-2001.
 XX 14-JUN-2001; 2001WO-US019444.
 XX 15-JUN-2000; 2000US-0212073P.
 XX 23-JUN-2000; 2000US-0213467P.
 XX 30-JUN-2000; 2000US-0215651P.
 XX 07-JUL-2000; 2000US-0216605P.
 XX 13-JUL-2000; 2000US-0218372P.
 XX 25-AUG-2000; 2000US-0228056P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
 PI Gandhi AR, Tribouley JW, Walla NK, Yao MG, Lu DAM, Greenwald SR;
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
 PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
 XX WPI: 2002-090207/12.
 XX N-PSDS; AAD26467.
 XX New polypeptides, useful for diagnosing, treating or preventing disorders
 PT of growth and development, cardiovascular and lipid, and diseases such as
 PT cancer, comprise human kinase polypeptides.
 XX Claim 1; Page 164-165; 197pp; English.
 XX The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for
 CC treating a disease or condition associated with decreased expression of

496 ELAPFARAKSLPPSPVTHSPLLLHPRGFLRPSASLPEEAASERSTEAPAPASPEGAGPP 555
 182 ELAPFARAKSLPPSPVTHSPLLLHPRGFLRPSASLPEEAASERSTEAPAPASPEGAGPP 241
 556 AAGCCVPRHVSIRSLFTHQAGESEPHALAPGSRHRHARRHLLKGVYIAGALPOLREPL 615
 242 AAGCCVPRHVSIRSLFTHQAGESEPHALAPGSRHRHARRHLLKGVYIAGALPOLREPL 301
 616 MEHRVLEEAAREEQATLLAKAPSFETALRLPASGTHLAPGSHSLHSDSPSTPSPSEA 675
 302 MEHRVLEEAAREEQATLLAKAPSFETALRLPASGTHLAPGSHSLHSDSPSTPSPSEA 361
 676 CGEAOQLPSAPSGAIPRDMGHPQGGKQIPSTGCHGHTAQPERRSPSPGOWAPAFCHPX 735
 362 CGEAOQLPSAPSGAIPRDMGHPQGGKQIPSTGCHGHTAQPERRSPSPGOWAPAFCHPX 421
 736 QGSAPOEGGCPHFAVAPCPGSPGSCKEAPLVPSSPFLGQAPAPAPAKASPELDSKM 795
 422 QGSAPOEGGCPHFAVAPCPGSPGSCKEAPLVPSSPFLGQAPAPAPAKASPELDSKM 481
 796 GPGDISLPGPKPGPCSSPGASQASASSQVSLRVGSSQVGTGPGPSLDAGTQEARDL 855
 482 GPGDISLPGPKPGPCSSPGASQASASSQVSLRVGSSQVGTGPGPSLDAGTQEARDL 541
 856 SDSTPTLQRPQEQVTKRFTSLGGRGGYAGVAGYGFAGGDDAGMGLGQGPWARIAMAVS 915
 542 SDSTPTLQRPQEQVTKRFTSLGGRGGYAGVAGYGFAGGDDAGMGLGQGPWARIAMAVS 601
 916 QSEEEQEEARASQSEEQEAEARAPLQVAGRPVPEVGRAPTRSSPEPTWEDIGQVS 975
 602 QSEEEQEEARASQSEEQEAEARAPLQVAGRPVPEVGRAPTRSSPEPTWEDIGQVS 661
 976 LVQIRDLSGDAEAADTISLIDSEVDPAVNLNLDYDIKYLPEFMI FRKPKSAQPPPS 1035
 662 LVQIRDLSGDAEAADTISLIDSEVDPAVNLNLDYDIKYLPEFMI FRKPKSAQPPPS 721
 1036 PMAEELAEFPETPTWPGELGPHAGLEITESESDVDALLAAAVGRKXKWSPPSRSLFH 1095
 722 PMAEELAEFPETPTWPGELGPHAGLEITESESDVDALLAAAVGRKXKWSPPSRSLFH 781
 1096 FPGRHLPDEPABGLRERVKASVEHISRLIKRPEGLEKEGPPRKPGLASFLSLGLKS 1155
 782 FPGRHLPDEPABGLRERVKASVEHISRLIKRPEGLEKEGPPRKPGLASFLSLGLKS 841
 1156 WDRAPTFLRELSDETIVLQGSVTLACQVSAQPAQAATWSKOGAPLESSSRVLI SATLKNF 1215
 842 WDRAPTFLRELSDETIVLQGSVTLACQVSAQPAQAATWSKOGAPLESSSRVLI SATLKNF 901
 1216 QLTILVVAEDLGVTCTSVNALGVTITGVLRKAERPSSPCPDICEVYADGVLLVWK 1275
 902 QLTILVVAEDLGVTCTSVNALGVTITGVLRKAERPSSPCPDICEVYADGVLLVWK 961
 1276 PVESYGPVTIVQCSLEGGSWTTLASDIPDCCYLTSKLSRGCTYTRTACVSKAGMGFYS 1335
 962 PVESYGPVTIVQCSLEGGSWTTLASDIPDCCYLTSKLSRGCTYTRTACVSKAGMGFYS 1021
 1336 SPSEQVLLGGPSHLASEESQGRSAQPLPSTKTAFQIQGRGSRVVRQCWKASGRAL 1395
 1022 SPSEQVLLGGPSHLASEESQGRSAQPLPSTKTAFQIQGRGSRVVRQCWKASGRAL 1081
 1396 AAKIIPYHPKDTAVLREYALKGRHPLAQLHAAYLSPRHLVLIILELCSGPPELLPCLA 1455
 1082 AAKIIPYHPKDTAVLREYALKGRHPLAQLHAAYLSPRHLVLIILELCSGPPELLPCLA 1141
 1456 ERASYSESEVKYLVQWMLGATVYLNQHLHLDLASENMITEYNLLKVDLGNQAQSLSQ 1515
 1142 ERASYSESEVKYLVQWMLGATVYLNQHLHLDLASENMITEYNLLKVDLGNQAQSLSQ 1201
 1516 EKVLPSSDKFKDYLETWAPLLEGGQAVPQTDIWAIGVTAFIMLSABYVPSSEGARDLQRG 1575
 1202 EKVLPSSDKFKDYLETWAPLLEGGQAVPQTDIWAIGVTAFIMLSABYVPSSEGARDLQRG 1261

PKIN and a composition comprising PKIN antagonist is useful for treating a disease or condition associated with overexpression of PKIN. The disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease, arteriosclerosis, anaemia, allergies, adult respiratory distress syndrome, autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease, osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's disease, rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis, bacterial, parasitic, fungal, viral, protozoal and helminthic infections) growth and development disorders (arteriosclerosis, cirrhosis, hepatitis, Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts), cardio vascular disease (arteriovenous fistula, hypertension, vasculitis, aneurysms, congestive heart failure, angina pectoris, myocarditis, ischaemic heart disease, chronic bronchitis, lung tumours); lipid disorder (fatty liver, Fabry's disease, Niemann-Pick's disease, hypcholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity of a test compound and in gene therapy. The present sequence is human PKIN-20 protein

Query Match	51.9%	Score	4533	DB 5	Length	871			
Best Local Similarity	100.0%	Pred. No.	2.8e-277						
Matches	871	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	795	MGPGDISLPGRKPGPCSSPGSASQASSQVSSLRVGSQVGTGPGSLDAEGWTOAED	854						
Db	1	MGPGDISLPGRKPGPCSSPGSASQASSQVSSLRVGSQVGTGPGSLDAEGWTOAED	60						
Qy	855	LSDSPTLQRPQEVTKRKSLSGRGGYAGVAGYGTFAFGDAGMQLGGPMWARIANAV	914						
Db	61	LSDSPTLQRPQEVTKRKSLSGRGGYAGVAGYGTFAFGDAGMQLGGPMWARIANAV	120						
Qy	915	SQSEEEQEERARAESQEEQEARAESPLPQVSARPVPEVGRAPTRSSPEPTPWEDIGQV	974						
Db	121	SQSEEEQEERARAESQEEQEARAESPLPQVSARPVPEVGRAPTRSSPEPTPWEDIGQV	180						
Qy	975	SLVQIRLSDGDAEAADTISLIDISEVDPAVNLSDLYDKYLPFFPMFKRVPKSAQPEPP	1034						
Db	191	SLVQIRLSDGDAEAADTISLIDISEVDPAVNLSDLYDKYLPFFPMFKRVPKSAQPEPP	240						
Qy	1035	SPMAEELABFPETPWPMPGELGPHAGLRITBESEDDVALLAEAAVGRKRWSPSRSLF	1094						
Db	241	SPMAEELABFPETPWPMPGELGPHAGLRITBESEDDVALLAEAAVGRKRWSPSRSLF	300						
Qy	1095	HFPGRHLPDEPABELGURERVKASVRHISRIILKGRPEGLEKEGPPRKKPGLASPRLSGLK	1154						
Db	301	HFPGRHLPDEPABELGURERVKASVEHISRIILKGRPEGLEKEGPPRKKPGLASPRLSGLK	360						
Qy	1155	SWDRAPTFRLRELSDETIVLQGSVTLACQVSAQPAQAQATWSKDGAPLESSSSRVLISATLKN	1214						
Db	361	SWDRAPTFRLRELSDETIVLQGSVTLACQVSAQPAQAQATWSKDGAPLESSSSRVLISATLKN	420						
Qy	1215	FQLLTILVVAEDLGVTVCVSNALGTVTITGVLRKAERPSSSPCPDGIQVYADGVLLVW	1274						
Db	421	FQLLTILVVAEDLGVTVCVSNALGTVTITGVLRKAERPSSSPCPDGIQVYADGVLLVW	480						
Qy	1275	KPVESYGPVTIYVQCSLEGGSWTTLASDIPDCCYLTSKLSRGYTVTRFATACVSKAGMPY	1334						
Db	481	KPVESYGPVTIYVQCSLEGGSWTTLASDIPDCCYLTSKLSRGYTVTRFATACVSKAGMPY	540						
Qy	1335	SSPSEQVLLGSPSHLASEESQGRSAQPLPSTKTFAQTQIRGRFVSVRQCKEASGRA	1394						
Db	541	SSPSEQVLLGSPSHLASEESQGRSAQPLPSTKTFAQTQIRGRFVSVRQCKEASGRA	600						
Qy	1395	LAAXIIPYHPKQKTAIVLREVEALKGRHPLHAQLHAAVLSPRHLVLIILEICSGPELLPCL	1454						
Db	601	LAAXIIPYHPKQKTAIVLREVEALKGRHPLHAQLHAAVLSPRHLVLIILEICSGPELLPCL	660						
Qy	1455	AERASYSSESEKYDWMQMSATQYLHNOHILHLDLRSENMITTEYNLLKVVYDLGNAQSLS	1514						

QY	1515	QEKVLSDKFKDYLETMAPELLEGGGAVPQTDIWAIGVTAFMLASBPVSSSEGARDLQR	1574
Db	721	QEKVLSDKFKDYLETMAPELLEGGGAVPQTDIWAIGVTAFMLASBPVSSSEGARDLQR	780
QY	1575	GLRKGLVRLSRVCYAGLSGGGAVAFRLSTLCAQFWRGPCASSCLOCFWLTTEGPGACSRPAPV	1634
Db	781	GLRKGLVRLSRVCYAGLSGGGAVAFRLSTLCAQFWRGPCASSCLOCFWLTTEGPGACSRPAPV	840
QY	1635	TFPTARLVFVRNREKRRALLYKRHNLAQVR	1665
Db	841	TFPTARLVFVRNREKRRALLYKRHNLAQVR	871
RESULT 10			
AAB30570			
ID	AAB30570 standard; protein; 871 AA.		
AC	AAB30570;		
XX	19-MAR-2001 (first entry)		
DE	A splice variant of a signal transduction polypeptide.		
XX	Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;		
KW	congestive heart failure; dilated congestive cardiomyopathy;		
KW	hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;		
KW	mitral valve disease; aortic valve disease; tricuspid valve disease;		
KW	myocardial infarction; cardiac arrhythmia; arteriosclerosis;		
KW	atherosclerosis; cardiac tumour; microbial infection; splice variant.		
XX	Homo sapiens.		
OS			
XX	WO2000063381-A1.		
FPN	26-OCT-2000.		
XX	PD		

11-APR-2000; 2000WO-US009488.
16-APR-1999; 99US-0129553P.
(SCIO-) SCIOS INC.
Zeng W, Stanton L, Kong H;
WPI; 2001-007013/01.
Novel h19G5 polypeptides capable of regulating signal transduction and
exhibiting kinase activity useful for identifying antibodies to treat
cardiac diseases, and additional mediators of signal transduction.
Claim 1; Page 74-76; 81pp; English.

The present sequence represents a splice variant of human α 1 signal transduction polypeptide. The polypeptide is designated H19G5. The protein is capable of regulating signal transduction and exhibits kinase activity. The H19G5 transcript is expressed in the heart. H19G5 polypeptides and polynucleotides are useful for preventing or treating a cardiac disease, such as congestive heart failure, dilated congestive cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve disease, aortic valve disease or tricuspid valve disease, angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary, arterial or rheovascular hypertension, arteriosclerosis, atherosclerosis and cardiac tumours in humans. The polypeptide is also useful for detecting the expression of a protein capable of regulating signal transduction or the expression of a protein capable of acting as a donor or acceptor molecule of a phosphate group. The monoclonal antibodies can be used as probes for detecting discrete antigens expressed by tissue or cell samples, and therefore used in humans for localization and monitoring of microbial infection.

Sequence 871 AA;

Query Match 51.8%; Score 4529; DB 4; Length 871;
Best Local Similarity 99.9%; Pred. No. 5e-277;
Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 795 MGPDISLGRPRGKPCSSQSSASQSSQSSVSLRVGSSQVTEGPPSLDAGWQTEARD 854
Db 1 MGPDISLGRPRGKPCSSQSSASQSSQSSVSLRVGSSQVTEGPPSLDAGWQTEARD 60

QY 855 LSDSTPTLQRPQVTKRFSLGGRGYAGVAGYGTFAFGDAGGMLGQGPWARIANAV 914
Db 61 LSDSTPTLQRPQVTKRFSLGGRGYAGVAGYGTFAFGDAGGMLGQGPWARIANAV 120

QY 915 SQSEEEHREARASQSEOEARASPLPOVSARVPVGVAPTRSPPEPTPREDIQV 974
Db 121 SQSEEEHREARASQSEOEARASPLPOVSARVPVGVAPTRSPPEPTPREDIQV 180

QY 975 SLVQIRDLSDAADAATISLDISEVDPAYLNSLDYDIKYLFFEFMIFRKVPKSAQPPPP 1034
Db 181 SLVQIRDLSDAADAATISLDISEVDPAYLNSLDYDIKYLFFEFMIFRKVPKSAQPPPP 240

QY 1035 SPVAEBELAEFPPTWPGELGPHAGLEITRESEVDALLAEAAVGRKWSPPSRSLF 1094
Db 241 SPVAEBELAEFPPTWPGELGPHAGLEITRESEVDALLAEAAVGRKWSPPSRSLF 300

QY 1095 HFPGRHLPLDEPAELGLRERVKASVEHISRIILKGRPEGLEKGPCKPGLASFLSLGLK 1154
Db 301 HFPGRHLPLDEPAELGLRERVKASVEHISRIILKGRPEGLEKGPCKPGLASFLSLGLK 360

QY 1155 SWDRAPELRELSDETAVLGQSVTLACQVSAQPAQATWSKDGAPLESSSVLISATLKN 1214
Db 361 SWDRAPELRELSDETAVLGQSVTLACQVSAQPAQATWSKDGAPLESSSVLISATLKN 420

QY 1215 FQLLTILVVAEDLGYVTVCSVSNALGTVTTTGLVRAERPPSSPCPDIGEVYADGVLLW 1274
Db 421 FQLLTILVVAEDLGYVTVCSVSNALGTVTTTGLVRAERPPSSPCPDIGEVYADGVLLW 480

QY 1275 KPVESGPTVYIYQCSLEGSSWTLASDIFDCYLTSLKSRGGTYTFTACVSKAGWGPY 1334
Db 481 KPVESGPTVYIYQCSLEGSSWTLASDIFDCYLTSLKSRGGTYTFTACVSKAGWGPY 540

QY 1335 SSPSEQVLLGGPGLHASEESQGRSAQPLPSTKTFAQTQIQGRFVSVRQWCKASGRA 1394
Db 541 SSPSEQVLLGGPGLHASEESQGRSAQPLPSTKTFAQTQIQGRFVSVRQWCKASGRA 600

QY 1395 LAAKITYHPKDTAVLREYALKGLRHPHLAQLAAVLSPHLVLIILELCSGPELLPCL 1454
Db 601 LAAKITYHPKDTAVLREYALKGLRHPHLAQLAAVLSPHLVLIILELCSGPELLPCL 660

QY 1455 AERASYSESEVXDYLQWMLSATQYLNQHLHLDRSENMIITEYNLLKVVYDLGNAQSL 1514
Db 661 AERASYSESEVXDYLQWMLSATQYLNQHLHLDRSENMIITEYNLLKVVYDLGNAQSL 720

QY 1515 QEKVLPDKFKDYLETWAPLELGGGAVPQTDIWAIGVTAFTMLSAEYPSVSEGAQLQR 1574
Db 721 QEKVLPDKFKDYLETWAPLELGGGAVPQTDIWAIGVTAFTMLSAEYPSVSEGAQLQR 780

QY 1575 GLKGLVRLSRVAGLGGGAVAFRLSTLCAQPWGRPCASSCLQCPWLTEEGFACSRPAP 1634
Db 781 GLKGLVRLSRVAGLGGGAVAFRLSTLCAQPWGRPCASSCLQCPWLTEEGFACSRPAP 840

QY 1635 TPEPTARLVFVRNREKRALLVYKRNLAQVR 1665
Db 841 TPEPTARLVFVRNREKRALLVYKRNLAQVR 871

RESULT 11
AAB30571
ID AAB30571 standard; protein; 548 AA.
XX
AC AAB30571;
XX
DT 19-MAR-2001 (first entry)

XX A full length rat signal transduction polypeptide.
DE Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
XX congestive heart failure; dilated congestive cardiomyopathy;
KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
KW mitral valve disease; aortic valve disease; tricuspid valve disease;
KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
KW atherosclerosis; cardiac tumour; microbial infection.
XX Rattus norvegicus.
OS
XX WO200063381-A1.
EN
XX 26-OCT-2000.
PD
XX 11-APR-2000; 2000WO-US009488.
PF
XX 16-APR-1999; 99US-0129553P.
PR
XX (SCIO-) SCIOS INC.
PA
XX Zeng W, Stanton L, Kong H;
FI WPI; 2001-007013/01.
XX
XX Novel H19G5 polypeptides capable of regulating signal transduction and
PT exhibiting kinase activity useful for identifying antibodies to treat
PT cardiac diseases, and additional mediators of signal transduction.
PS
XX Claim 1; Page 76-77; 81pp; English.
XX
XX The present sequence represents a rat signal transduction polypeptide,
CC designated H19G5. The human polypeptide is designated H19G5. The protein
CC is capable of regulating signal transduction and exhibits kinase
CC activity. The H19G5 transcript is expressed in the heart. H19G5
CC polypeptides and polynucleotides are useful for preventing or treating a
CC cardiac disease, such as congestive heart failure, dilated congestive
CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
CC mitral valve disease, aortic valve disease or tricuspid valve disease,
CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
CC arterial or rheovascular hypertension, arteriosclerosis, atherosclerosis
CC and cardiac tumours in humans. The polypeptide is also useful for
CC detecting the expression of a protein capable of regulating signal
CC transduction or the expression of a protein capable of acting as a donor
CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
CC be used as probes for detecting discrete antigens expressed by tissue or
CC cell samples, and therefore used in humans for localization and
CC monitoring of microbial infection
XX
XX Sequence 548 AA;

Query Match 28.0%; Score 2447.5; DB 4; Length 548;
Best Local Similarity 86.4%; Pred. No. 7.6e-146;
Matches 472; Conservative 27; Mismatches 46; Indels 1; Gaps 1;

QY 1121 HHSRIKGRPGLEKGPCKPGLASFLRSLGKSWDRAPTFRLSDETAVLGQSVTLA 1180
Db 3 HHSRIKGRPGLEKGPCKPGLASFLRSLGKSWDRAPTFRLSDETAVLGQSVTLA 62

QY 1181 CQVSAQPAQAQATWSKDGAPLESSSRVLSATLKNFQLLTILVVAEDLGVVTVCSVSNALG 1240
Db 63 CQVLAQPTAQTWSKDGALLESSGHLISLTKNFKQLLTILVVAEDLGVVTVCSVSNALG 122

QY 1241 TVTTTGVLRKAERPPSSPCPDIGEVYADGVLLWKPVESGPTVYIYQCSLEGSSWTLA 1300
Db 123 TAVTTGVLRKAERPPSSPCPDIGEVYADGVLLWKPVESGPTVYIYQCSLEGSSWTLA 162

QY 1301 SDIFDCCYLTSLKSRGGTYTFTACVSKAGMGPPSSPEQVLLGGPGLHASEESQGRS 1359
Db 183 SDISDCCYLTSLKSRGGTYTFTACVSKAGMGPPSSPEQVLLGGPGLHASEESQGRS 242

QY 1360 AQLPSTKTFAQTQIQGRFVSVRQWCKASGRAALAAKIIPHPKDTAVLREYALKG 1419

Db 243 AQLPSTKTFAQTQIRGRFSVVRQCRKASGRALAAKIYVQPEDKTVLREYALKR 302
 QY 1420 LRHPHLAQLHAAYLSPRHLVLILELCSGPPELLPCLAEASYSSESEVKDYLWMLSATQYL 1479
 Db 303 LHPHLAQLHAAYLSPRHLVLILELCSGPPELLPCLAEASYSSESEVKDYLWMLSATQYL 362
 QY 1480 HNOHILHLDRSENMIITEYNLLKVDLGNALSOEKVLPDCKPKDYLETWAPLLEGG 1539
 Db 363 HAQHILHLDRSENMMVTENLLKVIDLGNALSOEKVLPDCKPKDYLETWAPLLEGG 422
 QY 1540 GAVPQTDIWAIGVTAFLMISAEYFVSSEGARDLQGRKGLVRLSRVRCYAGLSGGAFAFLR 1599
 Db 423 GAVPQTDIWAIGVTAFLMISAEYFVSSEGARDLQGRKGLVRLSRVRCYAGLSGGAFAFLR 482
 QY 1600 STLCAQPMGRPCASSCLQCPMLTEEGPACSRPAPVTPPTARLRFVVRNKRKRALLYKRH 1659
 Db 483 SSLCAQPMGRPCASSCLQCPMLTEEGPACSRPAPVTPPTARLRFVVRNKRKRALLYKRH 1659
 QY 1660 NLAQVR 1665
 Db 543 NLAQVR 548

RESULT 12

AAB30572

ID AAB30572 standard; protein; 548 AA.

XX AC AAB30572;

XX DT 19-MAR-2001 (first entry)

XX DE A full length mouse signal transduction polypeptide.

XX KW Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
 KW congestive heart failure; dilated congestive cardiomyopathy;
 KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
 KW mitral valve disease; aortic valve disease; tricuspid valve disease;
 KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
 KW atherosclerosis; cardiac tumour; microbial infection.

XX OS Mus musculus.

XX PN WO200063381-A1.

XX PD 26-OCT-2000.

XX PF 11-APR-2000; 2000WO-US009488.

XX PR 16-APR-1999; 99US-0129553P.

XX PA (SCIO-) SCIOS INC.

XX PI Zeng W, Stanton L, Kong H;

XX PS WPI; 2001-007013/01.

XX PT Novel H19G5 polypeptides capable of regulating signal transduction and
 PT exhibiting kinase activity useful for identifying antibodies to treat
 PT cardiac diseases, and additional mediators of signal transduction.

XX PS Claim 1; Page 77-79; 81pp; English.

XX CC The present sequence represents a mouse signal transduction polypeptide,
 CC designated H19G5. The human polypeptide is designated H19G5. The protein
 CC is capable of regulating signal transduction and exhibits kinase
 CC activity. The H19G5 transcript is expressed in the heart. H19G5
 CC polypeptides and polynucleotides are useful for preventing or treating a
 CC cardiac disease, such as congestive heart failure, dilated congestive
 CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
 CC mitral valve disease, aortic valve disease or tricuspid valve disease,
 CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
 CC arterial or rheonovascular hypertension, arteriosclerosis, atherosclerosis

CC and cardiac tumours in humans. The polypeptide is also useful for
 CC detecting the expression of a protein capable of regulating signal
 CC transduction or the expression of a protein capable of acting as a donor
 CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
 CC be used as probes for detecting discrete antigens expressed by tissue or
 CC cell samples, and therefore used in humans for localization and
 CC monitoring of microbial infection

XX SQ Sequence 548 AA;

Query Match 27.6%; Score 2415.5; DS 4; Length 548;
 Best Local Similarity 85.7%; Pred. No. 8e-144;

Matches 468; Conservative 25; Mismatches 52; Indels 1; Gaps 1;

QY 1121 HISRILKGRPSGLKEGPRKPKGLASFRLSGLSKMDRAPTEIRLSDETVLQGSVTLA 1180
 Db 3 HISRILKGRPSGPRGPRKPKGLASFRLSGLSKMDRAPTEIRLSDETVLQGSVTLA 62

QY 1181 QVVSQAQAATWSKDGAPLESSSRVLIISATLKNFQLLTILVVVAEDLGVYTCVSNALG 1240
 Db 63 CQVLAQTAQATWSKDGVLLESSGHLIISLTKNFQLLTILVVKREDLGVYTCVSNPLG 122

QY 1241 TVTTTGVLRKAEKRPSSPCPDIGEVYADGVLLVWKPVESYGPVTYIVQCSLEGGSMTTLA 1300
 Db 123 TAVTTGVLRKAEKRPSSSPRPVEGELYKDAVLLVWKPVESGPTTYIVQCCIEGGSMTTLA 182

QY 1301 SDIPDCCVLTSLRGTYTFRACVSKAGNPGYSSPSEQVTLGGPSHLASEE-SQGRS 1359
 Db 183 SDISDCCVLTGKLRGGMYPFRACVSKAGNPGYSSPSEQVTLGGPSHLASEE-SQGRP 242

QY 1360 AQLPSTKTFAQTQIRGRFSVVRQCRKASGRALAAKIYPHPKDKTAVLREYALKR 1419
 Db 243 AQLPSTKTFAQTQIRGRFSVVRQCRKASGRALAAKIYPHPKDKTAVLREYALKR 302

QY 1420 LRHPHLAQLHAAYLSPRHLVLILELCSGPPELLPCLAEASYSSESEVKDYLWMLSATQYL 1479
 Db 303 LHPHLAQLHAAYLSPRHLVLILELCSGPPELLPCLAEASYSSESEVKDYLWMLSATQYL 362

QY 1480 HNOHILHLDRSENMIITEYNLLKVDLGNALSOEKVLPDCKPKDYLETWAPLLEGG 1539
 Db 363 HAQHILHLDRSENMMVTENLLKVIDLGNALSOEKVLPDCKPKDYLETWAPLLEGG 422

QY 1540 GAVPQTDIWAIGVTAFLMISAEYFVSSEGARDLQGRKGLVRLSRVRCYAGLSGGAFAFLR 1599
 Db 423 GAVPQTDIWAIGVTAFLMISAEYFVSSEGARDLQGRKGLVRLSRVRCYAGLSGGAFAFLR 482

QY 1600 STLCAQPMGRPCASSCLQCPMLTEEGPACSRPAPVTPPTARLRFVVRNKRKRALLYKRH 1659
 Db 483 SSLCAQPMGRPCASSCLQCPMLTEEGPACSRPAPVTPPTARLRFVVRNKRKRALLYKRH 1659

QY 1660 NLAQVR 1665
 Db 543 NLAQVR 548

RESULT 13

AAB65635

ID AAB65635 standard; protein; 2286 AA.

XX AC AAB65635;

XX DT 27-MAR-2001 (first entry)

XX DE Novel protein kinase, SEQ ID NO: 162.

XX KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiautomatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX OS Homo sapiens.

[illegible][illegible]

QY	1346	-----P H L S A E S E S G-----R S A Q P -L F S T-----	1366
Db	1838	V S F S P P T P P S Q A L S L K A V G P P P O T P P R R H R G L Q A A R P A B P T I P S T H V T P S E P K P F V L D	1897
QY	1367	-----	1366
Db	1898	T G T P I P A S T P Q G V K P V S S T V Y W T S F V S A P P A P E P P A P P P P P T K V T V Q S L S P A K E V	1957
QY	1367	-----K T F A F Q C I Q G R F S V V R C W E K A S G R A L A A K I P	1401
Db	1958	V S S P O S S P R S P R E G T I L R Q G P P Q P Y T F L E E K A R G F G V V R A C R E N A T G R T F V A K I V P	2017
QY	1402	Y H P K D T A V L R E Y E A L K G L R H P H L A Q L H A A Y L S P R H L V L I L E L C S G P E L L P C L A E R A S Y S	1461
Db	2018	Y A A E G K P R V Q E Y E V L A T L H E R I M S L H E A Y I T P S Y L V L I A E S C N R E L L C G L S D R F R Y S	2077
QY	1462	E S E V K D Y L W Q M L S A T O V L H N Q H I L H L D R S E N M I T E Y N L L K V D L G N A O S L S O E K V L P S	1521
Db	2078	E D D V A T Y W V Q L L G D Y L H G H V L H L D I K P D N L L A P D N A L K I V D F G S A P Y P N Q A L R E L	213
QY	1522	D K F R D Y L E T M A P E L L E C G A V P Q D T W A T G T A F I M L S A E Y P V S E G A R D O L G R L K G L V	1581
Db	2138	C H R T G T L E F M A P E N V M K G E P T G S A T D I W A G V I T Y I M L S R S F F Y E P D P Q E T A R I V G G R F	2197
QY	1582	R L S R C Y A G L S G G A V A F L R S T I C A Q P G R C A S S C L Q C P W L T E E G P A C S R P A P V T F T A R L	1641
Db	2198	D A F O L Y P N T S O S A T F L R K V L S V H P W S R P S L Q D C L A H P W L Q D A Y L M K L R Q T L T F T T N R L	2257
QY	1642	R V F V R N R K R A L L Y K H N L	1661
Db	2258	K E F L G E Q R R R A E A T H K V	2277
RESULT 14			
ID	AAE19160	standard; protein; 2380 AA.	
XX	AC	AAE19160;	
XX	DT	21-MAY-2002 (first entry)	
DE	DE	Human kinase polypeptide (PKIN-18).	
XX	XX	Human; kinase polypeptide; PKIN-18; gene therapy; Addison's disease; leukaemia; immune disorder; lymphoma; melanoma; developmental disorder; acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension; asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis; cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cystostatic; cancer; cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease; lipid disorder; fatty liver; Gaucher's disease; myocardial infarction; drug screening; transgenic animal; antiinflammatory; hepatotropic; hypotensive; anti-HIV; enzyme.	
OS	OS	Homo sapiens.	
FH	Key	Location/Qualifiers	
FT	Peptide	52..70	
FT	Protein	/label= Signal_peptide	
FT	Domain	71..2380	
FT	Domain	/note= "Mature human PKIN-18"	
FT	Domain	714..967	
FT	Domain	/note= "Eukaryotic proetin kinase domain"	
FT	Domain	2079..2331	
FT	Domain	/note= "Eukaryotic proetin kinase domain"	
PN	WO200208399-A2.		
XX	31-JAN-2002.		
PP	20-JUL-2001; 2001WO-US023092.		
XX	21-JUL-2000; 2000US-0220038P.		
PR	28-JUL-2000; 2000US-0222112P.		

PR	04-AUG-2000; 2000US-02282831P.	
PR	11-AUG-2000; 2000US-0224729P.	
XX	(INCY-) INCYTE GENOMICS INC.	
PA	(THOR/) THORNTON M.	
XX		
XX	Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK,	
PI	Patterson C, Ramkumar J, Gandhi AK, Policky JL, Baughn MK;	
PI	Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;	
PI	Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR;	
PI	Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;	
XX		
DR	WPI; 2002-206083/26.	
DR	N-PSDB; AAD30565.	
XX		
XX	New human kinase polypeptide, useful in diagnosis, prevention and	
PT	treatment of cancer, immune disorder, growth and developmental disorder,	
PT	cardiovascular disorder and lipid disorder.	
XX		
PS	Claim 1; Page 163-168; 196pp; English.	
XX		
CC	The present invention relates to an isolated human kinase polypeptide	
CC	(PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is	
CC	useful for diagnosing, treating and preventing cancer (e.g., leukemia,	
CC	lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency	
CC	syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's	
CC	disease, rheumatoid arthritis), a growth and developmental disorder (e.g.	
CC	burstis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a	
CC	cardiovascular disorder (e.g., atherosclerosis, hypertension), a	
CC	infarction), and a lipid disorder (e.g., fatty liver, cholestasis,	
CC	Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of	
CC	drug screening techniques and to analyse the proteome of a tissue or cell	
CC	type. PKIN is useful for creating knockin humanised animals or transgenic	
CC	animals to model human diseases, in somatic or germline gene therapy, to	
CC	generate a transcript image of a tissue or cell type, for detecting	
CC	differences in the chromosomal location due to translocation, inversion,	
CC	etc., among normal, carrier or affected individuals, and as hybridisation	
CC	probes for mapping naturally occurring genomic sequences. PKIN is useful	
CC	in southern or northern analysis, dot blot or other membrane-based	
CC	technologies, in PCR technologies, in dipstick, pin, multiformat enzyme	
CC	linked immunosorbent (ELISA)-like assays and in microarrays utilising	
CC	fluids or tissues from patients to detect altered PKIN expression. The	
CC	present sequence is human PKIN-18. Note: This sequence is said to be	
CC	encoded by PKIN-18 cDNA referred as SEQ ID NO:38 (AAD30565). However this	
CC	does not appear to be the case	
XX		
XX	Sequence 2380 AA;	
	Query Match 22.6%; Score 1973; DB 5; Length 2380;	
	Best Local Similarity 29.3%; Pred. No. 5.7e-115;	
	Matches 568; Conservative 236; Mismatches 638; Indels 498; Gaps 50;	
QY	54 RSMQVTTEDVCAQGTGTAPEALIEGDPOPSVTWVKDSQVLVDSTLSSQOEGTYSVLV 113	
Db	598 PAFPSIMDEVGAGETRRFVAVVVGKPLDPMWYKDEVLTESHVSFVVEECSLVV 657	
QY	114 RHVASKDAGVYTTCLAQNTGGQVLCKAELLVLGGDN----EPDSBKQSHR-RKLHSFPEVK 168	
Db	658 LSTGAQDGGVYTTCAQLAGEVSCKAELAVHSAQTAPEVEGVGDEDEHRRGRLLSDFYDIH 717	
QY	169 BEIGRGVFGFKRVQHKGNKILCAAKAFILRSTRQAAYRERDILAALSHPLVTGLDQF 228	
Db	718 QEIGKGAFSYLRRIVERSSSGLPEPAKFIPSOAKPKASARREARLARLQHDCVLYFHEAF 777	
QY	229 ETRKTLILILELCSBELDLRYKGVVTEARVKVYICQLVEGHYHLHSGVHLDTKPS 268	
Db	778 ERRGLVIVITELC-TEELLERIAKRTVCESIRAYMEQVLEGHYHLHSGVHLHDKPE 836	
QY	289 NILMVHPA--REDIKICDFGFAQNTTPAELOFSQYSGSPFVSPRIIOONPVSEASDIWAM 346	
Db	837 NLLVMDGAAGEQQRICDFGNQELTPPEQYVCQGTGTFEFVAPDINVQSPVSGVGTDTWV 896	
QY	347 GVISYLSLTCSPPFAGESDRATLLNVLEGRVSWSSPMAHLSEDAKQF-IKATLQAPQA 405	

Db 897 GWAFCLTGISPVGENDRTLLMIRNYNAFEETFLSLREARGFLIKVLQ--DRL 954
Qy 406 RPSAQCCLSHPWFLKSPAEAEHPIKQKFLAASRWORSILMSYKSLVMSSTPELIR 465
Db 955 RPTAEETLEHFWTKQAGAE---VSDHLKFLSRRWRQSOISYKCHLVKRPPELIR 1011
Qy 466 GPPDSPSLGVARHLCRTDGTSSSSSSSSDNEL-----APPARAK-SLPSPVTH 513
Db 1012 APPERVMTWTPRR--PPSPGGLSSSSSEELBELSPVRPLQPEPSGRSVLTDIPTED 1070
Qy 514 SPLHPRGFLRPSASLPEEAERSTAPAPASPEGAGPPAAQCVPRHVSIR----- 568
Db 1071 EALGTPTGATPMDWCEQRAFSQOEAFSPALFSPQEPAA--CASPRRELARGGSA 1129
Qy 569 -----SLFYHQAG--ESPEHGALAPG-----SRHPARRHLLK 600
Db 1130 ESALPRAGPRELGRGLKAAASVELPQRRSPGPGATLARGGLGEYEYQALQALRQLR 1189
Qy 601 GGYTAGALPGLREPLMEH-----RVLEBAAREEQATL-----LAKAPSPETALR 645
Db 1190 GGPEDGVKSLGRLPLESLGLGRARDPRMAAASSEAAHPHQPLENRGLQKSSFSQGEA 1249
Qy 646 LPASGTHLAPGHSLSLHDSFSTPR-----PSSACQEAQRLSPASGGAPIRDMGHPOGS 701
Db 1250 EP-RGRHRRAGAPLEIIVARLGARRLQESPSLSALSEAQ--PSSPA-----RPSAP 1297
Qy 702 KQLPSTGCHPQTAOPERPSDPSWGPQAPFCHPKQGSAPQEGCSHPAVAPCPGSPFG 761
Db 1298 K--PST--PKSAEFSATTFSDAPQPPAF--QAQDKAFEPPEPVRASKPAP-----PPQ 1346
Qy 762 SKC--EAPLVSPSPFLOQ-----POAPPAPAK-----ASPPDLKRMGP 797
Db 1347 ALQTLALPLTPYAQIIQSLQSLGHAQPSQGAAPPSEPKPAAPVAVASPP----- 1399
Qy 798 GDISLGRKPGKPCSPGSAQSSQSVSLRVSSQVGTETPCPSLDAEGWQEAEDLSD 857
Db 1400 -----PGAEKRVPSAGPPVLEAKRVPT-----VPPRFOSSLSSENLESEAVFE 1447
Qy 858 STPTLQRPQEQVTKRPSLGRGGYAGVAGYGTGTFAGDAGGMLGQGPMMARIAWAVSQS 917
Db 1448 A--KFKRSRES-----PLSLGLR-----LLSRS 1468
Qy 918 EEEQEERAR-AESQSEEQCARAESPLPQVSARVPVEGVRAPTRSPPEPTPWEDIGQVSL 976
Db 1469 RSEERGPGRGAEEDGYRSPAGTPL-----ELVRRPERS----- 1505
Qy 977 VQIRDLG--GDAEAADTISLDISEVDPAYLNLSDYDIKYLPEFPMIFRKVPKSAQPEPP 1034
Db 1506 --SVQDLRAVGEPLVRLSLSLSQ-----RLARTPPAQR----- 1538
Qy 1035 SPMAEEELAEFPEPTWPMGELGPHAGLETTERSEVDALLAEAAVGRKEKW----- 1086
Db 1539 -----HPANRARGDG-----ESSEGSARGSPVLAVRRRLSFTLERLS 1578
Qy 1087 -----SSFSRSLFHFGRIHPLDEPAELGRERVKASVEHI GRILKGRP----- 1130
Db 1579 SRLQSGSSSDSGASGRSTPL-----FGLRLRATSEGESLRL--GLPHNQLAAQAGAT 1631
Qy 1131 -----EGLEKEGP-----PRKKPGLASFRLSG-----LK 1154
Db 1632 TPSAESLGSEASATSGSAPGERSRLRWGFSRPRKDKGLSPNLSASVQEBLGHQYVRS 1691
Qy 1155 SWDRAPTFRLBELDETIVLQSVTLACQVSAQAPAAQATWSKQAPLESRRVLI SATLKN 1214
Db 1692 ESDFPFVPHIKLKDQVLEGEAATLLCLPAACAPAPI SWMKDKSLRSEPSVIIVSCDKG 1751
Qy 1215 FQLTLIVVAEDLGVTCTSVNALGTVTTTGVLRKAERPSSSPCCPDIGEVADGVLLVW 1274
Db 1752 RQLLSIPRAGKRHAGLYECATNVLGSIITSSCTVAVARVPGKLAPPEVPTQDTALVLM 1811
Qy 1275 KPVESYGVTVIVOCISLEGGG-WTTLASDIFDCCYLTSKLSRGCTTFTTACVSTAGMGP 1333

Db 1812 KPGRSAPCTYTILERRVDGESVMHPVSSGIPDCYYNVNTHLPVGVTVRPRVACANRAGQCP 1871
Qy 1334 YSSPSRQVLLGG----- 1345
Db 1872 FSNSEKVFVRGTQDSSAVPSAAHQEAPVTSRPARARPDSDTSLAPLAPAAPTPPSVT 1931
Qy 1346 -----PSHLASEESQG-----RSAQP-LPST----- 1366
Db 1932 VSPSSPPTSPSQALSSLKAVGPPQTPPRHRHGLQARPAEPTLPSHTVTFSEPKFFVLD 1991
Qy 1367 ----- 1366
Db 1992 TGTPIFASIPQGVKPVSSSTPVVVTSVFSAAPPAPPEPAPPPPTKVTVQSLSPAKEY 2051
Qy 1367 -----KTFAPQIOGRFVSVVQCKWCKASGRALAALPIIP 1401
Db 2052 VSSFGSPSSRSPRPECTTLRQGPQKPYTFLEKARGFVVRACKENATGTFVAKIYP 2111
Qy 1402 YHPKDKTAVLRVEALKGLRHPLAQLHAAVLSRHLVLIILELCSGPPELLPCLAEASYS 1461
Db 2112 YAAEGKRVLQYEVVLRTHHEEIMSLHEAYITPVLVLAESCGRNELLCCLSDFRYS 2171
Qy 1462 ESEVKDYLWQMSATQYLHNQIHLHLDURSENITEYNLLKVDLGNQAQSLSQSKVLPS 1521
Db 2172 EDDVATYVQLLQGLDYLGHGHVLDIKPDNLLAPDNALKIVDFGSAQFYNQALRPL 2231
Qy 1522 DKPKDYLEWMAPELLEGQAVPQTDIWAIGTAFIMLSAEYVPSSEGGARDLQGRKGLV 1581
Db 2232 GHRTGTLEFMAPEMKVGEPIGATDINGAGVLYIMLSGRSPFFYEPDQETEARIVGSRF 2291
Qy 1582 RLRSCTVAGLSGGNAVFLRSTLCAQWGRPCASSCLOCPLWTEEGPACSRPAPVPTFARL 1641
Db 2292 DAFQLYPNTSQSATFLRKVLSVHPMSRPSLQDCLAHPWLDQDAYLMKLRRLTFTTTNRL 2351
Qy 1642 RVFVRNRKRRALLVKRNL 1661
Db 2352 KEPLGQRRERAEATRHKV 2371
RESULT 15
ADE47676
ID ADE47676 standard; protein; 3208 AA.
XX AC ADE47676;
XX DT 29-JAN-2004 (first entry)
XX Human NOV14c protein SEQ ID NO:38.
DE human; cardiant; antiarteriosclerotic; hypotensive; immunosuppressive;
KW dermatological; anorectic; cytostatic; antidiabetic; haemostatic;
KW anti-HIV; antiasthmatic; antibacterial; virucide; neuroprotective;
KW neotropic; antiparkinsonian; antilipaemic; gene therapy; vaccine.
XX OS Homo sapiens.
XX WO2003076642-A2.
XX PN 18-SEP-2003.
XX PD 02-AUG-2002; 2002NO-US024459.
XX PF 02-AUG-2001; 2001US-0309501P.
XX PR 03-AUG-2001; 2001US-0310291P.
XX PR 08-AUG-2001; 2001US-0310951P.
XX PR 09-AUG-2001; 2001US-0311292P.
XX PR 13-AUG-2001; 2001US-0311979P.
XX PR 14-AUG-2001; 2001US-0312203P.
XX PR 17-AUG-2001; 2001US-0313156P.
XX PR 17-AUG-2001; 2001US-0313201P.
XX PR 20-AUG-2001; 2001US-0313702P.
XX PR 21-AUG-2001; 2001US-0314031P.
XX PR 23-AUG-2001; 2001US-0314466P.

PR 28-AUG-2001; 2001US-0315403P.
PR 29-AUG-2001; 2001US-0315853P.
PR 31-AUG-2001; 2001US-0315808P.
PR 21-SEP-2001; 2001US-0323936P.
PR 03-DEC-2001; 2001US-0338078P.
PR 05-FEB-2002; 2002US-0354555P.
PR 05-MAR-2002; 2002US-0361764P.
PR 13-APR-2002; 2002US-0373825P.
PR 15-MAY-2002; 2002US-0380971P.
PR 16-MAY-2002; 2002US-0380980P.
PR 28-MAY-2002; 2002US-0383761P.
PR 29-MAY-2002; 2002US-0383887P.
PR 01-AUG-2002; 2002US-00210130.
XX (CURA-) CURAGEN CORP.
XX Zernhusen BD, Patturajan M, Kekuda R, Miller CE, Rieger DK;
XX Pena CEA, Shinkens RA, Li L, Berghs C, Zhong M, Casman SJ, Voss EZ;
XX Boldog FI, Padigaru M, Smithson G, Shenoy SG, Ji W, Gorman L;
XX Vernet CAM, Leite MW, Guo X, Anderson DW, Spytek KA, Gerlach VL;
XX Burgess CE, Khramtsov NV, Ort T, Ellerman K, Rastelli L, Agee M;
XX Chaudhuri A, Chant JS, Dipippo VA, Edinger SR, Eisen A, Gangolli EA;
XX Giot L, Ooi CE, Rothenberg ME, Spaderna SK, Hjalte T, Liu X;
XX Taupier RJ, Catterton E;
XX WPI; 2003-779062/73.
XX N-PSDB; ADE47675.
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, diabetes,
XX atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
XX or pharmacogenomics.
XX Claim 1; SEQ ID NO 38; 562bp; English.
XX The invention relates to a novel (NOVX) human polypeptide. A polypeptide
XX of the invention has cardiac, antiarteriosclerotic, hypotensive,
XX immunosuppressive, dermatological, anorectic, cytotactic, antidiabetic,
XX haemostatic, anti-HIV, antitubercular, antibacterial, virocidic,
XX neuroprotective, nootropic, antiparkinsonian, and antilipidemic activity.
XX A polynucleotide encoding a polypeptide of the invention may have a use
XX in gene therapy, and as a vaccine. A polypeptide of the invention is
XX useful in the manufacture of a medicament for treating a syndrome
XX associated with a human disease, the disease selected from a pathology
XX associated with the polypeptide. These may also be used in diagnosis,
XX treating or preventing NOVX-associated disorders such as cardiomyopathy,
XX atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
XX haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
XX multiple sclerosis, infections, anorexia, cancer-associated cachexia,
XX neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
XX disease), haematopoietic disorders, dyslipidaemias and other wasting
XX disorders associated with chronic diseases. The nucleic acids are also
XX used as hybridisation probes, in chromosome mapping, tissue typing,
XX preventive medicine, and pharmacogenomics. The polypeptides are also
XX useful as vaccines. The present sequence represents a NOVX polypeptide of
XX the invention.
XX SQ Sequence 3208 AA;
Query Match 22.6%; Score 1973; DB 7; Length 3208;
Best Local Similarity 29.3%; Pred No. 8.6e-115;
Matches 568; Conservative 236; Mismatches 638; Indels 498; Gaps 50;
QY 54 PSMQVIERVQAGTGTQAFEMIEBGPQPSVYWKQSVQLVDSRSLSQOQGGTTLSVL 113
DB 1426 PFMESIMEVVGAGETARFAVVGKFLPDIMYKQEVLLTSSHSVYFYEENCSLV 1485
QY 114 RHVASKDAGVYVYCLAQNTGGQVLCRAELLVLGGDN---EPDSEKQSHR-RKLHSFYEVK 168
DB 1486 LSTGAQDGGVYVYCTAQNLAGEVSCAKELAVHSAQTAAMEVGVGDEDEHGRRLSDFYDIH 1545
QY 169 BEIGRGVFGFVRVQKGNKILCAKFIPLSRRTAQAYRERDILALSHPLVTGLLDQF 228

DB 1546 QEIGRGAFSYLRRIVERSSGLEFAAKFTPSQAKPKASARREARLARLHDCVLYFHEAF 1605
QY 229 ETRKTLIILILCSSEBELLDRLYKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLDKPS 288
DB 1606 ERRRGVIVITELC-TEELLERIAKFTVCESEIRAYMRQVLEGIHYLHSHGVHLDKPKE 1664
QY 289 NILMVHPA--REDIKICDFGFAQITPAELQFQSGSPFVSPFIIQQNPVSEASDIWAM 346
DB 1665 NLLVWDGAAGQVQRICDFGNAQLTGEQPCQYGTPEFVAPBIVNQSPVSGVTDIWPV 1724
QY 347 GVISYLSLTCSPFAGESDRATLNVLEGRVSWSPMAHLSEDAKDF-1KATLQAPQA 405
DB 1725 GVVAFCLTGTSPFVGENDRTLMNINYNVAFETTFLSRREARGLIKVLVQ--DRL 1782
QY 406 RPSAAQCLSHPWFLKSMPEAEAHFINTKQLKFLARSQRWQSLMSYKSLIMKRSIPELLR 465
DB 1783 RPTASETLEHFWFKTQAKGAE--VSTDHLKFLSRRRWQRQSIYSKCHLVLPIDELLR 1839
QY 466 GPDPSPSLGVARHLCDPFGSS 513
DB 1840 APPEKVVMTFRR--PPSGGGLUSSSSDEEBEELPSVPLQPFEPFSGSVLTDITED 1898
QY 514 SPLHPRGFLRPSASLPEAEASERSTEAPAPPASPEGAGPPAAQGCVPVRHSVIR---- 568
DB 1899 BALGTFETGAATPMDWQEQGRAPSDQEAPEALPSFGQEPAA--GASPRGELRRGSSA 1957
QY 569 -----SLFYHQAG--ESPFGALAPG-----SRHPARRRHLK 600
DB 1958 ESALPRAGPRELGLHKAASVELPQRRSPGFGNATRLARGLGEYGAQLQRLRLLR 2017
QY 601 GGYIAGALPGLREPLMEH-----RVLEEEAAREEQATL-----LAKAFSFETALR 645
DB 2018 GGPEDGVSGLRGLFLESLSLGGRRDPRMARAASSEAAPHQPLENRGLQKSSFSQCEA 2077
QY 646 LPASOTHLAPGHSHLSLEHDSPTPR-----PSSACGAQRLPSAPSGAFIRDMGHPOGS 701
DB 2078 EP-RGRHRRAGAPLEI PVARLGARLQESLSALSEAQ--PSPA-----RPSAP 2125
QY 702 KQLPSTGHPGTAQPERPSPDPMWGQAPFCHPXQSGAPOEGGSDHPAVACPFPGSPFP 761
DB 2126 K--PST---PKSAFSAFTTSPADAPQPPAP--QPAQDKAPRPERPVRASKAPAP--PPQ 2174
QY 762 SCK--EAPLVSPSSPFLQ-----POAPAPAK-----ASPPLDSKMG 797
DB 2175 ALQTLAULTYAGIIQLSLGHAQPSQGPAPSPSEPKFAHAFARVASPP----- 2227
QY 798 GDISLPGRPKPGPCSPGASQASSSQVSLRVGSSQVGTGPGSLDAGWTOEADLSD 857
DB 2228 -----PCAPEKRVPSAGGPPVLAFAKARVPT-----VPPRPGSSLSSTIENLESAVPE 2275
QY 858 STPTLQRPQEQVTRKFTSLGRRGGYAGVAGYGTFAFGDAGGMLGQGPMMARIAMAVSOS 917
DB 2276 A--XFKRSRES-----PLSLGLR-----LLSR 2296
QY 918 EEEQOEAR-AESQSEBQARAEPLPQVSAAPVPEVGRAPTRSPPEPTMEDICQVSL 976
DB 2297 RSEERGPFPGABEEDGIYRPSAGTPL-----ELVRRPERSR----- 2333
QY 977 VOIRDLG--GDAAADTISLDIGEDVAYLNLSDLYDIKYLPEFMIFRKVPKSAQPEP 1034
DB 2334 -SVQDLRAVGEPLVRLSLSLSQ-----RLRRTTPAQR----- 2366
QY 1035 SFMAEEELAEFPETWTPWPGELGPHAGLETSESDVALLAAVGRKRW----- 1086
DB 2367 -----HPAWEARGDGG-----ESSEGGSGARGSPVLAMRRRUSFTLERLS 2406
QY 1087 -----SGPSRLTFHPGRHLPLDEPAELGLRERVKASVEHISRLKGRP----- 1130
DB 2407 SRLQRSGSSSEDSGASGRSTPL-----FGRLRATSEGESLRL--GLPHNQLAAQAGAT 2459
QY 1131 -----EGLEKEGP-----PRKPGGLASPRLSG-----LK 1154

Db	2460	TPSABSLGSEASATSGSAPGESRSLRWGFSRPRKDKGLSPPNLSASVOEELGHQYVRS	2519
Qy	1155	SWDRAPFLRELSDETIVLQSVTLACQVSAQPAQAATWKGAPLESSESRVLISATLKN	1214
Db	2520	ESDPPVPHIKLQOVLEGEAATLLCLPAACPAHLSWMKDKSLASEPSVLIIVSCKOG	2579
Qy	1215	POLITILVVAEDIGVYTCVSNALGVTTTGVLRKAERSSSPCPDIGEVYADGVLLVW	1274
Db	2580	ROLLSIIPRAGKRHAGLYECSATNVLSITTSCTVAVARVPGKLAPPEVPQYQDTALVLW	2639
Qy	1275	KPVESYGPVTYIVQCSLEGGGS-WTTLASDIPDCCLYLSKLSRGCTYTFRTACVSKAGMGP	1333
Db	2640	KPGDSRAPCTYTLERRVDGESVWHPVSSGIPDCYVNVTHLPVGVTVRFRVACANRAGQGP	2699
Qy	1334	YSSPSEQVLLGG-----	1345
Db	2700	FSNSEKVFEVGTQDSSAVPSAAHQEAPVTSRPARARPPDSPTSLAPPLAPAAFTPSVT	2759
Qy	1346	-----PSHLASEESQG-----RSAQP-LPST-----	1366
Db	2760	VSPSSPPTPPSQALSSLKAVFPPTQTPRRHRGLQAARPAEPTLPSTHTVTPSEKPFVLD	2819
Qy	1367	-----	1366
Db	2820	TGTPIPASTPQGVKVPVSSSTPVVVVTSFVSAPPAPPEPPPEPTKVTVQSLSPAKEY	2879
Qy	1367	-----KTEAFOTQIORGRFSVVROCKEASGRALAAKIIP	1401
Db	2880	VSSPGSSPRSPREGTTLROGPQKPYTFLEEKARGFVVRACRENATGRFVAKIIP	2939
Qy	1402	YHPKDKTAVLREYKALKGRHPLHAQLHAAYLSRHLVLIILELCSGPELLPCLAEASYS	1461
Db	2940	YAAEGKRRVLQEVYVRLTHERIMSLHEAYITPRYLVLIAESCGNRELLCGLSDRFRYS	2999
Qy	1462	ESEVKDYLWQMLSATQYLNHONHILHDLRSENMLITEYNLLKVYDLGNAQSLSQEKVLP	1521
Db	3000	EDDVATYVQLLQGS-DYLLGHGHHVHLHDIKPNLLAPDNALKIVDFGSAQFYNPQALRPL	3059
Qy	1522	DKFKDYLETMAPELLEGGQVAPQIDWAIQVTAFLMSAEYFVSSEGARDLQRLKGLV	1581
Db	3060	GHRGTLEFMAPENVKGPISATDINGAGVLTIVMLSGRSFFYPDPQETEARIVGGRF	3119
Qy	1582	RLSRCYAGLSGGAVAFIRSTLCAQFWGRPCASSCLQCPWLTEEGPACSRPAPVTFPTARL	1641
Db	3120	DAFOLYPNTSQSATLFLRKVLVSHFWSRPSLQDCLAHFWLQDAYLMKLRQTLTFTTNR	3179
Qy	1642	RVFVNRKRRALLYKRHL 1661	
Db	3180	KEFLGEQRRRAEATRHKV 3199	

Search completed: April 23, 2004, 15:09:09
Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2004, 15:07:27 ; Search time 30 Seconds
(without alignments)
2865.241 Million cell updates/sec

Title: US-10-697-263-2

Perfect score: 8740

Sequence: 1 MGCCRLGCGGSAHVSQSG.....RNREKRALYKRNHNAQVR 1665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
- 5: /cgn2_6/ptodata/2/iaa/PCUTUS COMB.pap.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1665	4	US-09-858-664A-2
2	8740	100.0	1665	4	US-10-274-978-2
3	1173	13.4	846	4	US-09-858-664A-3
4	1173	13.4	846	4	US-10-274-978-4
5	893	10.2	549	4	US-09-858-664A-5
6	893	10.2	549	4	US-10-274-978-6
7	592.5	6.8	414	4	US-09-858-664A-13
8	592.5	6.8	414	4	US-10-274-978-14
9	591.5	6.8	2860	2	US-08-826-267-2
10	531	6.1	279	4	US-09-858-664A-4
11	531	6.1	279	4	US-10-274-978-5
12	521	6.0	298	4	US-09-858-664A-17
13	521	6.0	298	4	US-10-274-978-18
14	509.5	5.8	508	4	US-09-858-664A-18
15	509.5	5.8	508	4	US-10-274-978-19
16	504	5.8	260	4	US-09-858-664A-6
17	504	5.8	230	4	US-10-274-978-7
18	489	5.6	448	2	US-09-159-385-2
19	489	5.6	448	3	US-09-186-277-2
20	472.5	5.4	260	2	US-07-857-224B-23
21	466.5	5.3	454	2	US-09-159-385-1
22	466.5	5.3	454	3	US-09-186-277-1
23	466	5.3	274	4	US-09-858-664A-14
24	466	5.3	274	4	US-10-274-978-15
25	465	5.3	358	4	US-09-230-896C-29
26	463.5	5.3	1050	4	US-09-428-711A-16
27	456	5.2	331	3	US-08-810-712-24

ALIGNMENTS

RESULT 1

US-09-858-664A-2
; Sequence 2, Application US/09858664A
; Patent No. 6482624

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui, et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL000927-CIP

; CURRENT APPLICATION NUMBER: US/09/858,664A

; CURRENT FILING DATE: 2001-05-17

; PRIOR APPLICATION NUMBER: 09/711,134

; PRIOR FILING DATE: 2000-11-11

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 1665

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-858-664A-2

Query Match 100.0%; Score 8740; DB 4; Length 1665;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCCRLGCGGSAHVSQSGLTNHPSPVGGWHPGLCGWGGHSSLPALPGPPSMQVTI 60

DB 1 MGCCRLGCGGSAHVSQSGLTNHPSPVGGWHPGLCGWGGHSSLPALPGPPSMQVTI 60

QY 61 EDVQAGTGTGTAQFEAIIIEGDPQPSVTWYKDSVQLVDSVTRLSQQQEGTYSLVLRHVASKD 120

DB 61 EDVQAGTGTGTAQFEAIIIEGDPQPSVTWYKDSVQLVDSVTRLSQQQEGTYSLVLRHVASKD 120

QY 121 AGVYTCIAQNTGGGVLCXAELLVGGNEPDSKQSHRKLHSHFYKVEIGRGVGFVK 180

DB 121 AGVYTCIAQNTGGGVLCXAELLVGGNEPDSKQSHRKLHSHFYKVEIGRGVGFVK 180

QY 181 RVQHGKNIKCAAFIPURSTRQAQYRERDILAAALSHPLVTGLLOQFETRKTLIILEL 240

DB 181 RVQHGKNIKCAAFIPURSTRQAQYRERDILAAALSHPLVTGLLOQFETRKTLIILEL 240

QY 241 CSSEBLLDLRYKGVVTEAEVKVYIQQLVEGLHYLHSHGVLDIKPSNLMWHPAREDI 300

DB 241 CSSEBLLDLRYKGVVTEAEVKVYIQQLVEGLHYLHSHGVLDIKPSNLMWHPAREDI 300

QY 301 KICDFGFAQNTTRAELOFQSGSPFVSPEIIQONPVSEASDIWAMGVISYLSLTCSPP 360

DB 301 KICDFGFAQNTTRAELOFQSGSPFVSPEIIQONPVSEASDIWAMGVISYLSLTCSPP 360

QY 361 AGESDRATLNLVLEGRVSWSPMAAHLSEDAKFIKATLQAPQAPRPSAAQCLSHPWFLK 420
Db 361 AGESDRATLNLVLEGRVSWSPMAAHLSEDAKFIKATLQAPQAPRPSAAQCLSHPWFLK 420
QY 421 SMPAEAAHFTNTKQLFLLARSQRSLMSYSKILVMSIPELLRGPPDPSLGVARHLK 480
Db 421 SMPAEAAHFTNTKQLFLLARSQRSLMSYSKILVMSIPELLRGPPDPSLGVARHLK 480
QY 481 RDTGGSSSSSSSSONELAPPAKSLPPSPVTHSPHLLHPRGFLRPSASLPEAEASERST 540
Db 481 RDTGGSSSSSSSSONELAPPAKSLPPSPVTHSPHLLHPRGFLRPSASLPEAEASERST 540
QY 541 EAPAPPASPEGAGPPAAQGVCPVHRSVIRSLFVHQAQESPEHGALAPGRRRHPRARHLK 600
Db 541 EAPAPPASPEGAGPPAAQGVCPVHRSVIRSLFVHQAQESPEHGALAPGRRRHPRARHLK 600
QY 601 GGYIAGALPGLRPLMEHRVLEBAABEQAOTLLAKAPFETALRLPASGTHLAPCHSHS 660
Db 601 GGYIAGALPGLRPLMEHRVLEBAABEQAOTLLAKAPFETALRLPASGTHLAPCHSHS 660
QY 661 LEHDSPTSPSPSEACGEAQLRPSAPSGGAPIRDMGHPOGSKQLPSTGHPCTAQPERPS 720
Db 661 LEHDSPTSPSPSEACGEAQLRPSAPSGGAPIRDMGHPOGSKQLPSTGHPCTAQPERPS 720
QY 721 PDSFPGQAPFCHPKQGSAPQECSCSPHAPVAPCPGSPPPGSCCKEAPLVPSPFLGQQA 780
Db 721 PDSFPGQAPFCHPKQGSAPQECSCSPHAPVAPCPGSPPPGSCCKEAPLVPSPFLGQQA 780
QY 781 PPAPAKASPLDKNMGPGDITSLPGRKPGPCSPSPGASQASSQSSVSLRVSSQVGTBP 840
Db 781 PPAPAKASPLDKNMGPGDITSLPGRKPGPCSPSPGASQASSQSSVSLRVSSQVGTBP 840
QY 841 PSIDAEGWTQAEADLSDSTPTLQRPQOVTMRKFSLGRRGAGVAGTGTAFGGDAGM 900
Db 841 PSIDAEGWTQAEADLSDSTPTLQRPQOVTMRKFSLGRRGAGVAGTGTAFGGDAGM 900
QY 901 LGQGPMMARAWAVSQSEEEQEAARAEQSEBEQEAARAEPLQVSPARPVEVGRAPTR 960
Db 901 LGQGPMMARAWAVSQSEEEQEAARAEQSEBEQEAARAEPLQVSPARPVEVGRAPTR 960
QY 961 SSPEPTWEDIGQVSLVQIRDLSDGDAEADTISLISEVDPAAYINSLDIKYLPEEFM 1020
Db 961 SSPEPTWEDIGQVSLVQIRDLSDGDAEADTISLISEVDPAAYINSLDIKYLPEEFM 1020
QY 1021 IFRKVPKSAQPEPPSPMAEEELAEFPEPTWPGELGPHAGLEITEESEDVDALLAAV 1080
Db 1021 IFRKVPKSAQPEPPSPMAEEELAEFPEPTWPGELGPHAGLEITEESEDVDALLAAV 1080
QY 1081 GRKHKWSSPSRSLFHPGRHLPLDEPAELGLRERVKASVEHSIRILKGRPELKEGPPR 1140
Db 1081 GRKHKWSSPSRSLFHPGRHLPLDEPAELGLRERVKASVEHSIRILKGRPELKEGPPR 1140
QY 1141 KKPGLASFRLSGKSWDRAPTFRLRLSDTETVLGQSVTLACQVSAQPAQAATWSKDGAPL 1200
Db 1141 KKPGLASFRLSGKSWDRAPTFRLRLSDTETVLGQSVTLACQVSAQPAQAATWSKDGAPL 1200
QY 1201 ESSSRVLIATLKNFOLLTLVVAEDLGVTCSVSNALGTVTITGVLRKARSSSPCP 1260
Db 1201 ESSSRVLIATLKNFOLLTLVVAEDLGVTCSVSNALGTVTITGVLRKARSSSPCP 1260
QY 1261 DIGEYVADGVLVYKVPESGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGGTYT 1320
Db 1261 DIGEYVADGVLVYKVPESGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGGTYT 1320
QY 1321 PRATCVSKAGVYSSPSEOVLLGGPSHLASEEESQGRSAQPLPSTKTFATQIQGRGF 1380
Db 1321 PRATCVSKAGVYSSPSEOVLLGGPSHLASEEESQGRSAQPLPSTKTFATQIQGRGF 1380
QY 1381 SVVRQWKEKASGRALAAKIIYPHKDKTAVLREYALKGLRPHPLAQLFAAYLSPRHLVL 1440
Db 1381 SVVRQWKEKASGRALAAKIIYPHKDKTAVLREYALKGLRPHPLAQLFAAYLSPRHLVL 1440
QY 1441 ILELCSPPELLPCLAEASVSESEVKDYLWQMLSATQYLLHQHILHLDLRSENMIITEYN 1500

Db 1441 ILELCSPPELLPCLAEASVSESEVKDYLWQMLSATQYLLHQHILHLDLRSENMIITEYN 1500
QY 1501 LLKVVLDLNGAQSLSQEKVLPSPDKFYLETMAPELLEGGAVPOTDIWAGVTAFIMLSA 1560
Db 1501 LLKVVLDLNGAQSLSQEKVLPSPDKFYLETMAPELLEGGAVPOTDIWAGVTAFIMLSA 1560
QY 1561 EYPVSESGARDLQRLKGLVRLSRCYAGLSGGAVAFRLSTLCQAPWGRPCASSCLOCPW 1620
Db 1561 EYPVSESGARDLQRLKGLVRLSRCYAGLSGGAVAFRLSTLCQAPWGRPCASSCLOCPW 1620
QY 1621 LVEEGPACSRPAPVTFPTARLRFVVRNREKRALLYKHNLAQVR 1665
Db 1621 LVEEGPACSRPAPVTFPTARLRFVVRNREKRALLYKHNLAQVR 1665

RESULT 2
US-10-274-978-2
; Sequence 2, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274, 978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858, 664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711, 134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Human
US-10-274-978-2

Query Match 100.0%; Score 8740; DB 4; Length 1665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCCLRLGCGGCSVAHVSQGLTNHPSVMYCGWHFPGLCGWWGGLHSSLPALPGPSPMQVTI 60
Db 1 MGCCLRLGCGGCSVAHVSQGLTNHPSVMYCGWHFPGLCGWWGGLHSSLPALPGPSPMQVTI 60
QY 61 EDVQATGCTAQTAFQAEIIEGDPQPSVTWYKDSVQLVDSTLSQQOEGTYSYLVLRHVASKD 120
Db 61 EDVQATGCTAQTAFQAEIIEGDPQPSVTWYKDSVQLVDSTLSQQOEGTYSYLVLRHVASKD 120
QY 121 AGVYTCLAQNTGGQVLCKAELLVGLGDNPEPSEKQSHRRKLHSPYEVKEEIGRGVGFVK 180
Db 121 AGVYTCLAQNTGGQVLCKAELLVGLGDNPEPSEKQSHRRKLHSPYEVKEEIGRGVGFVK 180
QY 181 RVQHKGNKILCAKEFIPLSRTRAQAYREDDILAAALSHPLVTGLDDOPETKTLILILEL 240
Db 181 RVQHKGNKILCAKEFIPLSRTRAQAYREDDILAAALSHPLVTGLDDOPETKTLILILEL 240
QY 241 CSSEELLDELRYKGVVTEAEVKVYIQQVLEGLHYLHSHGVHLHLDIKPSNIMLVHPAREDI 300
Db 241 CSSEELLDELRYKGVVTEAEVKVYIQQVLEGLHYLHSHGVHLHLDIKPSNIMLVHPAREDI 300
QY 301 KICDGFAGNITPAELQSQSGSEFEVSPETIIQQNPVSEASDIWAMGVISYLSLTCSSPF 360
Db 301 KICDGFAGNITPAELQSQSGSEFEVSPETIIQQNPVSEASDIWAMGVISYLSLTCSSPF 360
QY 361 AGESDRATLNLVLEGRVSWSPMAAHLSEDAKFIKATLQAPQAPRPSAAQCLSHPWFLK 420
Db 361 AGESDRATLNLVLEGRVSWSPMAAHLSEDAKFIKATLQAPQAPRPSAAQCLSHPWFLK 420
QY 421 SMPAEAAHFTNTKQLFLLARSQRSLMSYSKILVMSIPELLRGPPDPSLGVARHLK 480

Db 421 SYPAEAEHFINTKQLKELLARSQRSLMSYKSLVWRSIFPELLRGPPDPSLGVAREHL 480
Qy 481 RDTGSSSSSSSSNEIAPFARAKSLPPSPVTHSFLHPRGFURPSASLPEEAEASERST 540
Db 481 RDTGSSSSSSSSNEIAPFARAKSLPPSPVTHSFLHPRGFURPSASLPEEAEASERST 540
Qy 541 EAPAPPASPEGAGPAAQCGVPRHSVIRSLFYHOAGESPHEGALAPGSRHRPARRHLK 600
Db 541 EAPAPPASPEGAGPAAQCGVPRHSVIRSLFYHOAGESPHEGALAPGSRHRPARRHLK 600
Qy 601 GGYIAGALPGLREPLMEHRVLEEEAAEEQATLLAKAPSPETALRLPASGTHLAPGHSHS 660
Db 601 GGYIAGALPGLREPLMEHRVLEEEAAEEQATLLAKAPSPETALRLPASGTHLAPGHSHS 660
Qy 661 LEHDSPTSPRESSACCEAOLPAPSGGAPIRDMGHPQSGKOLPSTGGHPGTAQPERPS 720
Db 661 LEHDSPTSPRESSACCEAOLPAPSGGAPIRDMGHPQSGKOLPSTGGHPGTAQPERPS 720
Qy 721 PDSWGPAPFCHPKQSGAPQEGSPHAPVPCPPGSPFGSCKEAPLVPSSPFLGQPOA 780
Db 721 PDSWGPAPFCHPKQSGAPQEGSPHAPVPCPPGSPFGSCKEAPLVPSSPFLGQPOA 780
Qy 781 PPAKASAPPLDSKMGPDIDSLPGRPKGPCSSPGSASQASSSSQVSLRVGSSQVGTPEG 840
Db 781 PPAKASAPPLDSKMGPDIDSLPGRPKGPCSSPGSASQASSSSQVSLRVGSSQVGTPEG 840
Qy 841 PSLDAEGHTQBAEDLSDSTPLQRPQEQVTRKFTSLGGRGYAGVAGYGFAGGADAGM 900
Db 841 PSLDAEGHTQBAEDLSDSTPLQRPQEQVTRKFTSLGGRGYAGVAGYGFAGGADAGM 900
Qy 901 LQCGPMWARIWAVSQSEEEQEARAESQSEEQEARAESPLPQVSARPVPEVGRAPTR 960
Db 901 LQCGPMWARIWAVSQSEEEQEARAESQSEEQEARAESPLPQVSARPVPEVGRAPTR 960
Qy 961 SSPEPTWEDIGQVSLVOIRDLSGDAEAADTISLDISEVDAYLNLSLDYDKLPEEFM 1020
Db 961 SSPEPTWEDIGQVSLVOIRDLSGDAEAADTISLDISEVDAYLNLSLDYDKLPEEFM 1020
Qy 1021 IPRKVPKSAQPEPPSPMAEEELAEPEPTWMPGELPHAGLEITESESDVALLAAAV 1080
Db 1021 IPRKVPKSAQPEPPSPMAEEELAEPEPTWMPGELPHAGLEITESESDVALLAAAV 1080
Qy 1081 GRKRWSSPSSLPHFPGRHPLDPEAPGLRERVKASVEHISILKRPGLKEGPPR 1140
Db 1081 GRKRWSSPSSLPHFPGRHPLDPEAPGLRERVKASVEHISILKRPGLKEGPPR 1140
Qy 1141 KXPGLASFLSGLKSWDRAPTFRLSDVTVVVGOSVTLACQVSAQAAQATWSKDGAPL 1200
Db 1141 KXPGLASFLSGLKSWDRAPTFRLSDVTVVVGOSVTLACQVSAQAAQATWSKDGAPL 1200
Qy 1201 ESSSRVLISATIKNFQLLTILVVAEDIGVYTCVSNALGTVTTGVLKKAERSSSPCP 1260
Db 1201 ESSSRVLISATIKNFQLLTILVVAEDIGVYTCVSNALGTVTTGVLKKAERSSSPCP 1260
Qy 1261 DIGEVVADGVLLVMKPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGSTYT 1320
Db 1261 DIGEVVADGVLLVMKPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGSTYT 1320
Qy 1321 FRTACVSKAGMGFSPSPSEQVLLGSPSHLASEESQGRSAQPLPSTKTFAFQTOIQRGF 1380
Db 1321 FRTACVSKAGMGFSPSPSEQVLLGSPSHLASEESQGRSAQPLPSTKTFAFQTOIQRGF 1380
Qy 1381 SVVROCKEASGRALAAKTIIPYHPKDKTAVLREYALKGLRHPHQAQLHAAYLSPRHVL 1440
Db 1381 SVVROCKEASGRALAAKTIIPYHPKDKTAVLREYALKGLRHPHQAQLHAAYLSPRHVL 1440
Qy 1441 ILELCSGPELLPCLABRASYSSEVKDYLMWLSATQYLNHQLHLDRSENMIITEYN 1500
Db 1441 ILELCSGPELLPCLABRASYSSEVKDYLMWLSATQYLNHQLHLDRSENMIITEYN 1500
Qy 1501 LLKYVDLGNASLSQEKVLPSPDKFDYLTMAPELLEGQVAPQTDIWAIGVTAIFMLSA 1560

Db 1501 LLKYVDLGNASLSQEKVLPSPDKFDYLTMAPELLEGQVAPQTDIWAIGVTAIFMLSA 1560
Qy 1561 EYPVSEGNARDLQGRKGLVRLSCYAGLSGGAVALRSTLCAQPGWGRPCASSCLQCPW 1620
Db 1561 EYPVSEGNARDLQGRKGLVRLSCYAGLSGGAVALRSTLCAQPGWGRPCASSCLQCPW 1620
Qy 1621 LTEEGPACSRPAPVTFPTARLRFVVRNREKRALLYKRNLAQVR 1665
Db 1621 LTEEGPACSRPAPVTFPTARLRFVVRNREKRALLYKRNLAQVR 1665
RESULT 3
US-09-858-664A-3
; Sequence 3, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-3
Query Match 13.4%; Score 1173; DB 4; Length 846;
Best Local Similarity 34.9%; Pred. No. 5.7e-62;
Matches 313; Conservative 121; Mismatches 316; Indels 146; Gaps 29;
Qy 54 PSMQVTEVDQAQTGTAQFAEAIIEGDPQPSVTWYKDSVQLVDSTRLSQOQGTGTVSLV 113
Db 1 PRFESIMEDVEVGAGETARFVAVVEGKPLPDIMWYKDEVLLTSSHSVFWYENECSLV 60
Qy 114 RHVASDKAGVYTCLAQNTGQVLCRAELLVLGDN---EPDSEKQSHR-RKLHSFEVVK 168
Db 61 LSTGAQDGGVYTCTAQNLAGVSCAEALVHSAQTAAMEVEGVEDDHGRRLUSDYDILH 120
Qy 169 ERTGRGVFGVVRVQKGNKILCAAKFIPLRSTRQAAYRERDILAAALSHPLVTGLDQF 228
Db 121 QETGRGAFSYLRIVERSSGLEFAKFIPOAKPKASARREARLLARLQHDCLVYFHEAF 180
Qy 229 ETKTULILELCSBELLDRLYRKGVVTEAEVKVIQQLVEGLVHLHSHGVHLHDIKPS 288
Db 181 ERRRGVIVITELC-TEELLERARKPTVCESTRATMRQVLEGIHVHSHVHLHDVKEP 239
Qy 289 NIIMVHPA--REDIKICDFGAQNTPAELOFSQVSPFVSPFVPEIIQQNPVSEASDIWAM 346
Db 240 NLVWDGAGQEQVRCIDFGNAQELIPGEPQCOYGTPEFAPEIVNQSPVSGVTDIPV 299
Qy 347 GVTSYLSLTCSSPFAGESDRATLLNLYLEGRVSWSSPMAHLSEDADKF-IKATLQAPQA 405
Db 300 GVVAFLCLAGISPFVGENDRRTLLMNIRYNVAPEETTFLSLREARGFLIKVLQ--DRL 357
Qy 406 RPSAAQCLSHPWFLKMPAEAEHFINTKQLKELLARSQRSLMSYKSLVWRSIPELLR 465
Db 358 RPTAEETLEHPWFKTQAKGAB---VSTDHLKFLSRRRWQSRQISYKCHLVLPPELLR 414
Qy 466 GPDPSPSLGVARHLCDTGGSSSSSSSSDNEL-----APFARAK-SLPPSPVTH 513
Db 415 APFVVMVTPRR-PPSGGLSSSSSEEELEELPSVPRPLQPEFSGRSVSLTDIPTED 473
Qy 514 SPLHPRGFLRPSASLPEEAEASERSTAPAPASPEGAGPPAAQCVPRHSVIR----- 568
Db 474 EALGTPTGAATPMQEQRAPSDQEQAPSPALPSQOEPAAGASPARGLRRGSSA 532

569 -----SLFVHQAG--ESPHEGALAPG-----SRHPARRHLK 600
533 ESALPRAGPRELGRGLHKAASVELPQRRSPGGATRLARGGEGEYACQALQRLR 592
601 GGYIAGALPGLRPLMEH-----RVLEEEAAREEQATL-----LAKAPSPETALR 645
593 GGPEDGKVSGLRGLPGLLESIGGRARDPRMARAAASSEAAAPHQPLENRLGKSSSFSQGEA 652
646 LPASGTHLAPGHSHLSHDSPTPR-----PSSEACGEAORLPSAPSGGAPIRDGMHPOGS 701
653 EP-RGRHRRAGAPLETPVARLGARLQESPSLSALSEAQ--PSSPA-----RPSAP 700
702 KQLPSTGGHGTGAQPERPSPDPMGQAPFCHPKQGSAPQEGSCSPHPAVAPCPGSPFP 761
701 K--PST---PKSAEPSATTPSDAQPAP--QPAQDKAPEPRPEVRASKAP-----PPQ 749
762 SCK--EAPLVPSPPFLQG-----POAPPAPAK-----ASPPLDSKMP 797
750 ALQTLALPLTPYAQIIQSLQSHAGQPSQGPAAPSEPKPAAVFAVAP-----802
798 GDISLPGRPKPGPCSPGASQASSQVSSLRVGSQVTEPGPSLDAGWQTQAB 853
803 -----PGAPEKRVPSAGGPPVLAEKARVPT-----VPPRPGSSLSSENLESE 846

RESULT 4
US-10-274-978-4
; Sequence 4, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Human
US-10-274-978-4

Query Match 13.4%; Score 1173; DB 4; Length 846;
Best Local Similarity 34.9%; Pred. No. 5.7e-62;
Matches 313; Conservative 121; Mismatches 316; Indels 146; Gaps 23;

54 PSMQVTEVQAGTGAQTAQFAIIEGDPQPSVTWYKDSVQLVDSVTRLSQQQEGTYSVLV 113
1 PPFESIMEDVEVGAGETARFAVVEGKPLDMMYKDEVLLTSSSHVSVFVNEECSLVV 60
114 RHVASKDAGVYTCLAQNTGGVLCVLAELVLGSDN-----EPDSEKQSHR-RKLHSFYEVK 168
61 LSTCAQDGGVYTCATNLAGVSCKALVAHSAQTAAMEVGEVGEDEHGRRLSDFFDIH 120
169 EETGRGVGVKRVQHKNGKILCAKFIPLRSTRQAQYRERDILAAHSHPLVTGLLDQF 228
121 QETGRGAFSVLRIVERSSGLSEFAAKFIPSOAKPKASARREARLLARLQHDVLYFHEAF 180
229 ETRKTIILILELCSSELDRLRYKGVVTEAAVKVYIQLVGEGLYHSHGVLDIDKPS 288
181 ERRGLVIVTELC-TSELLERARKPTVCEIRAYMQVLEGIHYLQSHVHLVDKPE 239
289 NILMVTAPA--REDIKICDFGPAQNTTPAELQFSQVSGSPFVSPEIQQNVFSEASDIWAM 346
240 NLLVMDGAAGEQQVRCIDFGNAQELTPGEPQCYQGTPEFVAPEVNGSPVSGVTDIWPV 299

347 GVISVLSLTCSPFFAGESDRATILNVLEGRVSWSPMAHILSEDAKOP-IKATIQAPQA 405
300 GVAVFLCLTGTSFFVGENDRITLMNRYNVAFFETTFSLSRREARGFLIKVLVQ--DRL 357
406 RPSAAQCLSHPWFLKMPAEBAHFINTKQLKFLARSQWQSLMSYKSIILVMRSTPELLR 465
358 RPTAEETLEHPWFKTQAKGAE--VSTDHLKFLSRRWQSQISKYKHLVLRIPPELLR 414
466 GPPDPSLGVARHLCDTGGSSSSSSSDNEL-----APFAK-SLPSPSVTH 513
415 APPERVWVMPRR-PPPGGLSSSSDSEEELELPSVPRPLQEFSGSRVSLTDITD 473
514 SPILLHFRGLRPSASLPSAEAEASERSTAPAPPASPEGAGPPAAQCGVPRHSVIR-----568
474 EALGTPTGAATPMQWQGRAPSDQQAEPSEALPSPGQEPAA-GASPRRGLRGRSSA 532
569 -----SLFVHQAG--ESPHEGALAPG-----SRHPARRHLK 600
533 ESALPRAGPRELGRGLHKAASVELPQRRSPGGATRLARGGEGEYACQALQRLR 592
601 GGYIAGALPGLRPLMEH-----RVLEEEAAREEQATL-----LAKAPSPETALR 645
593 GGPEDGKVSGLRGLPGLLESIGGRARDPRMARAAASSEAAAPHQPLENRLGKSSSFSQGEA 652
646 LPASGTHLAPGHSHLSHDSPTPR-----PSSEACGEAORLPSAPSGGAPIRDGMHPOGS 701
653 EP-RGRHRRAGAPLETPVARLGARLQESPSLSALSEAQ--PSSPA-----RPSAP 700
702 KQLPSTGGHGTGAQPERPSPDPMGQAPFCHPKQGSAPQEGSCSPHPAVAPCPGSPFP 761
701 K--PST---PKSAEPSATTPSDAQPAP--QPAQDKAPEPRPEVRASKAP-----PPQ 749
762 SCK--EAPLVPSPPFLQG-----POAPPAPAK-----ASPPLDSKMP 797
750 ALQTLALPLTPYAQIIQSLQSHAGQPSQGPAAPSEPKPAAVFAVAP-----802
798 GDISLPGRPKPGPCSPGASQASSQVSSLRVGSQVTEPGPSLDAGWQTQAB 853
803 -----PGAPEKRVPSAGGPPVLAEKARVPT-----VPPRPGSSLSSENLESE 846

RESULT 5
US-09-858-664A-5
; Sequence 5, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-5

Query Match 10.2%; Score 893; DB 4; Length 549;
Best Local Similarity 44.0%; Pred. No. 1.7e-45;
Matches 179; Conservative 79; Mismatches 135; Indels 14; Gaps 7;

54 PSMQVTEVQAGTGAQTAQFAIIEGDPQPSVTWYKDSVQLVDSVTRLSQQQEGTYSVLV 113
147 PPFESIMEDVEVGAGETARFAVVEGKPLDMMYKDEVLLTSSSHVSVFVNEECSLVV 206
114 RHVASKDAGVYTCLAQNTGGVLCVLAELVLGSDN-----EPDSEKQSHR-RKLHSFYEVK 168

Db 207 LSTGAQGGVYCTAONLAGEVSKAEALAVHSAQTAAMEVEGVGEDEHRRRLSDFYDIH 266
QY 169 BEIGRGVFGVKRVQHKGNKILCAAKFIPLRSRTRACAVRERDILAAALSHPLVTGLDQF 228
Db 267 QEIGRGAFSYLRIVERSSGLEFAKFIPLSQAKPKASARREALLARLQHDVCVLYFHEAF 326
QY 229 ETRKTLILILELCSBELDLRYKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKPS 288
Db 327 ERRRLVITVTELC-TELLERIAKPTVCESERAYMRQVLEGIHYLHSHGVHLHDVKPE 385
QY 289 NILMVHPA--REDIKICDFGAQNTITPAELQFSGYSGSPFVSPEIIOQNPVSEASDIWAM 346
Db 386 NLLVWDGAGQVQVRIQDFGNAQELTFGEPOYCYGTPTPEFVAPEIYNQSPVSGVTDIWPV 445
QY 347 GVISYLSLTCSSPAGESDRATLLNVLEGRVSWSSPMAHLSEDAKDF-IKATLQAPQA 405
Db 446 GVVAFLCLTGISPPFVGENDRITLNMIRYNVAPEETFTLSLSREARGFLIKVLVQ--DRL 503
QY 406 RPSAAQCLSHFWFLKSPAEAEAFINTKQLKFLARSRWORSLSYK 452
Db 504 RPTAETLEHPWPKTOAKGAE--VSTDHLKLFLLRRRWORSLSYK 547
RESULT 6
US-10-274-978-6
; Sequence 6, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274, 978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Human
US-10-274-978-6
Query Match 10.2%; Score 893; DB 4; Length 549;
Best Local Similarity 44.0%; Pred. No. 1.7e-45;
Matches 179; Conservative 79; Mismatches 135; Indels 14; Gaps 7;
QY 54 PSMQVTTIEDVQAQTGGTAQFEAIIIEGDPQSPVTVYKDSYQVLDSTRLSQQGGTYSVLV 113
Db 147 PRFSSIMEDVEVGAGETARAVVVEGKPLDINWYKDEVLLETSHSVFVVEENCSLVV 206
QY 114 RHVASKDAGVYTCIAQNTGGQVLCKAELLVLGDN----EPSEKQSHR-RKLHSFYEVK 168
Db 207 LSTGAQGGVYCTAONLAGEVSKAEALAVHSAQTAAMEVEGVGEDEHRRRLSDFYDIH 266
QY 169 BEIGRGVFGVKRVQHKGNKILCAAKFIPLRSRTRACAVRERDILAAALSHPLVTGLDQF 228
Db 267 QEIGRGAFSYLRIVERSSGLEFAKFIPLSQAKPKASARREALLARLQHDVCVLYFHEAF 326
QY 229 ETRKTLILILELCSBELDLRYKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKPS 288
Db 327 ERRRLVITVTELC-TELLERIAKPTVCESERAYMRQVLEGIHYLHSHGVHLHDVKPE 385
QY 289 NILMVHPA--REDIKICDFGAQNTITPAELQFSGYSGSPFVSPEIIOQNPVSEASDIWAM 346
Db 386 NLLVWDGAGQVQVRIQDFGNAQELTFGEPOYCYGTPTPEFVAPEIYNQSPVSGVTDIWPV 445
QY 347 GVISYLSLTCSSPAGESDRATLLNVLEGRVSWSSPMAHLSEDAKDF-IKATLQAPQA 405
Db 446 GVVAFLCLTGISPPFVGENDRITLNMIRYNVAPEETFTLSLSREARGFLIKVLVQ--DRL 503
QY 406 RPSAAQCLSHFWFLKSPAEAEAFINTKQLKFLARSRWORSLSYK 452
Db 504 RPTAETLEHPWPKTOAKGAE--VSTDHLKLFLLRRRWORSLSYK 547

Db 446 GVVAFLCLTGISPPFVGENDRITLNMIRYNVAPEETFTLSLSREARGFLIKVLVQ--DRL 503
QY 406 RPSAAQCLSHFWFLKSPAEAEAFINTKQLKFLARSRWORSLSYK 452
Db 504 RPTAETLEHPWPKTOAKGAE--VSTDHLKLFLLRRRWORSLSYK 547
RESULT 7
US-09-858-664A-13
; Sequence 13, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-13
Query Match 6.8%; Score 592.5; DB 4; Length 414;
Best Local Similarity 34.5%; Pred. No. 1e-27;
Matches 145; Conservative 69; Mismatches 191; Indels 15; Gaps 6;
QY 53 PSMQVTTIEDVQAQTGGTAQFEAIIIEGDPQSPVTVYKDSYQVLDSTRLSQQGGT 107
Db 1 PPEFVPLSEVTCETGETVVLRCVCRGPKASITWKGPHEHTLNNDGHVYSISYDLGEAT 60
QY 108 TYSVLVHVASDAGVYTCIAQNTGGQVLCKAELLVLGDNPDSEKSHRRKLHSPYEV 167
Db 61 ---LKVGVTTEDDGIYTCIAVNDMGSSASSASLRYLGP--MDGIMVTKDNFDSFYSE 115
QY 168 KEBIGRGVFGVKRVQHKGNKILCAAKFIPLRSRTRACAVRERDILAAALSHPLVTGLDQ 227
Db 116 VAEIAGRGFSVWKKCDQKGTKEAVATKFNKKLMKRDQVTHLGILOSLQHEPLLVLGDDT 175
QY 228 PETRKTLLILELCSSEELLRLYKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKP 287
Db 176 FETPTSYILLVLEMAQDQGRLLDCVWRWGLTEGKIRAHLEVEAVRYLHNCRIAHLDKP 235
QY 288 SNILMVHP-AREDIKICDFGAQNTITPAELQFSGYSGSPFVSPEIIOQNPVSEASDIWAM 346
Db 236 ENILVDESIAKPTIKLADFGDAVQNTTYYIHQLLGNFEFAAPEIILGNPVSLTSDTWSV 295
QY 347 GVISYLSLTCSSPAGESDRATLLNVLEGRVSWSSPMAHLSEDAKDFIKATLQAPQA 406
Db 296 GVITYLLSGVSGFFLDSDVEETCLNICRLDPSFPDDYFKGVSKAKBFVCFLLQEDPAK 355
QY 407 PSAAQCLSHFWFLKSPAEAEAFINTKQLKFLARSRWQ---RSLMSYKSLVWMSIPEL 463
Db 356 PSAAALAQEQW-LQAGNRSTGVLDTSRLTSFIERRKHNDVPRISIKNLFQSLPLPRV 414
RESULT 8
US-10-274-978-14
; Sequence 14, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274, 978

ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragoras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2860 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-826-267-2

Query Match 6.8%; Score 591.5; DB 2; Length 2860;
Best Local Similarity 34.7%; Pred. No. 1.4e-26;
Matches 145; Conservative 68; Mismatches 190; Indels 15; Gaps 6;

QY 53 PPSQVTTIEDVQAGTGTGTAQFEALIEGDPQPSVTVYKDSVQLVD-----STRLSQQQEGT 107
DB 1 PPFVPLSEVTCETGETVLCRCVGRPKASITWKGEHNTLNDGHYSISYDLGEAT 60
QY 108 TYSVLRLHVASKDAGVYVTCIAQNTGGVLCCKAEILLVGLGDNEDPSEKSHRRKLHSHFYEV 167
DB 61 ---LKIVGVTTEDDGIYTCIAVNDMGSSASSASLRVLGPG--MDGIMVTVKDNFDSFYSE 115
QY 168 KEBIGRGVGFVKRVQHKGNKILCAAKFIPLRSETRAQAYRERDITLALSHPLVTVGLDQ 227
DB 116 VAEIHRGRSVVKKCDQKTKAVATKFNKKLMKRDQVTHELGILQSLQHPVLVGLDQ 175
QY 228 FETPKTLILILELCSSEELLDRLYRGVGVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKP 287
DB 176 FETPTSYILVLEMDQGRLLDCVVRWGSLTEGKIRAHLGVELEAVRYLHNCRIAHLDLKP 235
QY 288 SNILMHP-AREDIKICDFGFAQNTITPAELQFSQVSPFVSPELQONPVSEASDIWAM 346
DB 236 ENILVDESIAKPTIKLADFGDAVOLNTYYIHQLGNPEFAPEIILGNPVSLTSDTWSV 295
QY 347 GVYSLSLTCSPFAGESDRATILNVLEGRVSWSSPMAHLSEDAKDFIKATLQAPQAR 406
DB 296 GVITYVLLSGVSPFLDSDVEETCLNICRLDFFSPDDYFKGVSKAKFVCFLLQEDPAKR 355
QY 407 PSAAQCLSHPWFLKMPAEAEHFNKQLKFLARSWQ---RSLMSYKSLVWMSIP 463
DB 356 PSALALQEQW-LOAGNGRSTGVLDTSRLTSFIERRKHQNDVPIRSIKNFKLSRLLP 414

RESULT 9
US-08-826-267-2
Sequence 2, Application US/08826267
Patent No. 5994070
GENERAL INFORMATION:
APPLICANT: Streuli, Michel
TITLE OF INVENTION: No. 5994070el TRIC Molecules and Uses Related Thereto
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/826,267
FILING DATE: 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/014,214
FILING DATE: 27 MARCH (1996)

ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragoras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2860 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-826-267-2

Query Match 6.8%; Score 591.5; DB 2; Length 2860;
Best Local Similarity 34.7%; Pred. No. 1.4e-26;
Matches 145; Conservative 68; Mismatches 190; Indels 15; Gaps 6;

QY 53 PPSQVTTIEDVQAGTGTGTAQFEALIEGDPQPSVTVYKDSVQLVD-----STRLSQQQEGT 107
DB 2448 PPFVPLSEVTCETGETVLCRCVGRPKASITWKGEHNTLNDGHYSISYDLGEAT 2507
QY 108 TYSVLRLHVASKDAGVYVTCIAQNTGGVLCCKAEILLVGLGDNEDPSEKSHRRKLHSHFYEV 167
DB 2508 ---LKIVGVTTEDDGIYTCIAVNDMGSSASSASLRVLGPG--MDGIMVTVKDNFDSFYSE 2562
QY 168 KEBIGRGVGFVKRVQHKGNKILCAAKFIPLRSETRAQAYRERDITLALSHPLVTVGLDQ 227
DB 2563 VAEIHRGRSVVKKCDQKTKAVATKFNKKLMKRDQVTHELGILQSLQHPVLVGLDQ 2622
QY 228 FETPKTLILILELCSSEELLDRLYRGVGVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKP 287
DB 2623 FETPTSYILVLEMDQGRLLDCVVRWGSLTEGKIRAHLGVELEAVRYLHNCRIAHLDLKP 2682
QY 298 SNILMHP-AREDIKICDFGFAQNTITPAELQFSQVSPFVSPELQONPVSEASDIWAM 346
DB 2683 ENILVDESIAKPTIKLADFGDAVOLNTYYIHQLGNPEFAPEIILGNPVSLTSDTWSV 2742
QY 347 GVYSLSLTCSPFAGESDRATILNVLEGRVSWSSPMAHLSEDAKDFIKATLQAPQAR 406
DB 2743 GVITYVLLSGVSPFLDSDVEETCLNICRLDFFSPDDYFKGVSKAKFVCFLLQEDPAKR 2802
QY 407 PSAAQCLSHPWFLKMPAEAEHFNKQLKFLARSWQ---RSLMSYKSLVWMSIP 461
DB 2803 PSALALQEQW-LOAGNGRSTGVLDTSRLTSFIERRKHQNDVPIRSIKNFKLSRLLP 2859

RESULT 10
US-09-858-664A-4
Sequence 4, Application US/09858664A
Patent No. 6482624
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES THEREOF
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL000927-CIP
CURRENT APPLICATION NUMBER: US/09/858,664A
CURRENT FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 09/711,134
PRIOR FILING DATE: 2000-11-11
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 279
TYPE: PRT
ORGANISM: Homo sapiens
US-09-858-664A-4

Query Match 6.1%; Score 531; DB 4; Length 279;
Best Local Similarity 39.5%; Pred. No. 2.8e-24;
Matches 111; Conservative 45; Mismatches 123; Indels 2; Gaps 1;

Qy	1336	SPSEQVLLGGPSHLASBESOGSRAQPLSTKTFAPQTQIQGRFFSVVROCWEKASGRAL	1395
Db	1	SPAKEYSSPGSGSPRPGTTLRGQPKQPFYLEEKARGFGVVRACRENATGRTF	60
Qy	1396	AAKTIYPHKDKTAVLREYEALKGLRPHLAQLHAAALSPRHVLILELCSGELLPLCLA	1455
Db	61	VAKIVPAAEGKPVLOEYEVRLTLHERIMSLHEAVITPRYLVLAESCGNRELLCGLS	120
Qy	1456	ERASYSESVKDYLMQWLSATQVYLHNOHILHLDRSENMLITENYLKKVVDLGNAGLSLQ	1515
Db	121	DRFYSEDDVATYVQLLQGLDYLHGHHVLHLDIKPNLLLAPDNALKIVDFSGAPQYNP	180
Qy	1516	EKVLPSDKFDYLTMTAPLLEGGGAVQPQTDIWAIGVATFIMLSAEYFVSSEGARDLQRG	1575
Db	181	QALRPLGHRGTGLTFMAEMVKGEPIGSATDIMGAGVLYTMLSGRSPFFYEPDPQETEAR	240
Qy	1576	LRKGLVRLSRCYAGLSGGAVAFRLSTLCAOPGWRPCASSCL	1616
Db	241	IVGGRFDFAFOIYPNTSOSATIFRKVLVSHVWAPP--SSCL	279

```

RESULT 11
US-10-274-978-5
; Sequence 5, Application US/10274978
; Patent NO. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00927-CIP-DIV
; CURRENT FILING DATE: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Human
; US-10-274-978-5

```

Query Match	6.1%;	Score 531;	DB 4;	Length 279;
Best Local Similarity	39.5%;	Pred. No. 2.8e-24;		
Matches 111;	Conservative 45;	Mismatches 123;	Indels 2;	Gaps 1;
Qy	1336	SPSEQVLLGSPSHLASBEEQSGRSAPLPSTKTFAFQTIQGRGFSVVRQCEKASGRAL	1395	
Db	1	SPAKEVVSFGSSPSRPSRPEGTTLRQGPFPQKPYTFLEEKARGRGFGVRAACRENATGRTP	60	
Qy	1396	AAKIIPYHPKDKTAVLREYEALKGLRHPLAAQLHAAVLSPRHLVLTILECSGPELPCIA	1455	
Db	61	VAKIIPVAAGKPKRVLOGEYVLRTHLHRTMSLHEAIIIPRYLVLTAAESCGNRELLCGLS	120	
Qy	1456	ERASYSSSEVKDYLMQMSLATQYLHNQHILHLDLRSENMIITBYNLLKVVLDLNAQSLQ	1515	
Db	121	DRFRYSDDVATYMQVLLQGLDLVGHGHVLLHLDIKPNLLLAPDLNAKIVDFGSAQFPNP	180	
Qy	1516	EKVLPSDKFDXYLETMAPELLEGQAVPQDDIWAIGVTAIFMLSAEVPSVSEGARDLQRG	1575	
Db	181	QALRPLGHRGTCTLEFPAMENVKGEPIGSATDINGAGVLTVMLSGRSPFVEPPDPQTEAR	240	
Qy	1576	LRKGLVRLSCYAGLSGGAFAIIRSTICAQFWRPCASSCL	1616	
Db	241	IVGGRFADFQIYPNNTQSATLFLRKVLSVHPWGRP--SSCL	279	

RESULT 12
US-09-858-664A-17

```

; Sequence 17, Application US/09858664A
; Patent No. 6482824
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; THEREOF
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1000927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-858-664A-17

```

Query Match	5.0%;	Score 521;	DB 4;	Length 298;
Best Local Similarity	35.2%;	Pred. No. 1.2e-23;		
Matches 105;	Conservative	69;	Mismatches 123;	Indels 2; Gaps 2;
QY	159	RKLHSPYEVKEBIGRGVFGFKRVQHGKNKILCAAKFPIPLRSRTRAQAYR-ERDILAALS	217	
DB	1	QKVSDFYDIERLGSRGFGGVFLVEKTKTRKWAGKFFKAYSAKENIRQEIISIMNCLH	60	
QY	218	HPLVTGILDQFENRKTLLILILELSCSSBEELDRLYKGV-VTEABKVYVIOQLVEGLHYLH	276	
DB	61	HPLKVQCDAFEKAKNIWMVLEIVSGSELPERIDEDFELTERECIKVMRQISGVEYIH	120	
QY	277	SHGVHLIDTPSNILMWHYPAREDIKICDFGAQNITPAELQFSQYGSPFVSPILIQNP	336	
DB	121	QKGTVHLLDPENIMCYNKTGTIKLDFGLGARLENAGSLKVLFGTFEFVAPVINYEP	180	
QY	337	VSEASDIWAGVTSYLSLTCSPPFAGSDSDRATILNVLEGRVSSPMAAHLSEDAKFIK	396	
DB	181	ISYATDMSIGVICYLLVSGLSFPMGNDNETLANVTSATWDFDDEAFDISDDAKDFIS	240	
QY	397	ATLQAPCAPSAOCLSHMFLKSPFAEAAHINTKQLFELLARSWOSRLMSYKSI	454	
DB	241	NLLKKMNKRLDCTCQLQHPWLMKDTKMBAKKLSKDRMKKYKARRKQWGHAVRAI	298	

```

RESULT 13
US-10-274-978-18
; Sequence 18, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Human
US-10-274-978-18

```

Query Match	6.0%;	Score 521;	DB 4;	Length 298;
Best Local Similarity	35.2%;	Pred. No. 1.2e-23;		
Match 195 - Conservation	68.	Mismatched 122.		
			Models	2.
			Cans	2.

Query Match	5.8%;	Score	509.5;	DB	4;	Length	508;
Best Local Similarity	25.5%;	Pred. No.	1.2e-22;				
Matches	125;	Conservative	75;	Mismatches	173;	Indels	117;
Gaps	5;						
QY	48	PALFPGPMSQVTTIEDVQAQTGGTAQFAEIIIEGDPQPSVTWYKDSVQLVDSTRLSQOQEGT	107				
Db	18	PKAAMPQIIQFPEDQKVRAGESVELFGKVTGQPICTCTWMPKPKQIDSEHIKVNSEN	77				
QY	108	TYSILVLRHVASKAGAVTTCIAQTGGVOVLCKAELLVLG-----	145				
Db	78	GSKLTILAAQEHCGCYLLVENKLGSRQAVNLTVVDKDPDPAGTTCASDIRSSSLTLS	137				
QY	146	-----	146				
Db	138	WYGSYDGGSAVQSYSIEIWDSSANKTWKELATCRSTSNFVQDLLDFHEYKFRVRAINVG	197				
QY	147	DNEP-----DSEKQ-----SHRRKLHSFYEVKEEIG	172				
Db	198	TSEPSQSESLTVGKEPPEPKMKWRCCQTDDKEPEVDYRTVTINTEQKVSDFYDIERLG	257				
QY	173	RGVFGFKRVQHKNGKILCAAKFIPLRSRTQAQYR-ERDILAAISHPLVTGLLDQFETR	231				
Db	258	SGKEQGVFLRVKTKRWAGCKFFKAYSAKENIRQEISITMNCILHFKPLVQCVDAAFEK	317				
QY	232	KTLTILILELCSBELLDRLYRKGV-VTEAEVKVYIQQLVEGLHYLHSHGVLLHLDIKPSNI	290				
Db	318	ANIWMVLEIVSGGELFERIIDEDFELTETRECIKWMRQISEGVEYTHKGIVHLDLKPENI	377				
QY	291	LMWHFAPREDIKCDFGAQNTTPAELQPSQYSGSPFVSPETIIQNPPVSEASDIWAMGVIS	350				

Query Match	5.8%;	Score	509.5;	DB	4;	Length	508;
Best Local Similarity	25.3%;	Fred. No.	1.2e-22;	Indels	117;	Gaps	5;
Matches	125;	Conservative	75;	Mismatches	173;		
QY	48	PALPGPPSQVTTIEDVQAQTGGTAQAEATIEGDPQPSVTWYKDSVLQDSTRJSSQQEGT	107				
DB	18	PXAAMPPIQIPEDQKVRAGESVELFGVTGTPITCTWMPKFKQIQDSEHIKVNSEN	77				
QY	108	TYSLVLRHVASKDAGVYTCLAQNTGGQVLCKABELVLG-----	145				
DB	78	GSKLTILAAQEHCGCYTLLVENKLGSRQAVNLTVVDKPDPPAGTFCASDIRSSSLTLS	137				
QY	146	-----G	146				
DB	138	WYSSVDGGGSAVOSYISIEIWDSSANKTWKELATCRSTSFNVQDLLPDHEYKFRVRAINVG	197				
QY	147	DNPF-----DSEKQ-----SHRKLSHFYEVKEIG	172				
DB	198	TSEFQSQSELTUVGEKPEEPKMKWRQOTDDEKSEFVDYRTVTTNTEQKVSDFYDIEERLG	257				
QY	173	RGVFGFKVRQVHKGNKILCAAKEIPLRSRTRAQAYR-ERDILAAALSHPLVTGLDQFPETR	231				
DB	258	SGHGGQVFLVEKTKRWAGKFFKAYSACEKINRQEI SIMNCLHHPKLVQCVDAPEEK	317				
QY	232	KTILILELCSBELDLRLVRKV- VTEAEVKYIQQLVEGLHVLHSHGVHLHDIKESNI	290				
DB	318	ANIMVLEIVSGELPERIIDEFELTERECIKYMRQISEGVEYTHQGVHLHDKESNI	377				
QY	291	LMVHPAREDIKIDFGFAQNIITPAELQFSYGSPEFVSPBIIQONPVSEASDIWAMGVIS	350				
DB	378	MCVNKTGTR-KLIDFGLARLENAWSLKVLFGEFVAPVINYEPISYATDMWSIGVIC	437				
QY	351	YLSUTCSSPFAGSDRATLNLVLEGRVSWSSPMAAHLSDSAKOFIKATIQAPQAPRSA	410				
DB	438	YILVSGLSFPWGDNDNETLANVTSATWDFDDEAFDEISDDAKOFISNLLKKOMKNRLDCT	497				
QY	411	QCLSHPWFLK	420				

Db 498 QCLQHPWLMK 507

Search completed: April 23, 2004, 15:12:06
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2004, 15:11:23 / Search time 59 Seconds
(without alignments)
7802.224 Million cell updates/sec

Title: US-10-697-263-2

Perfect score: 8740

Sequence: 1 MGCCRLGCGCVSAHVSQGLTNHPSVWGWHPLGCGGGLHSSLPALPPPSMQVTI 1665

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1133595 seqs, 276475211 residues

Total number of hits satisfying chosen parameters: 1133595

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1665	9	US-09-858-664A-2
2	8740	100.0	1665	12	US-10-697-263-2
3	8727	99.9	1665	12	US-10-415-011-22
4	8423	96.4	2630	13	US-10-077-130-2
5	8423	96.4	7968	13	US-10-077-130-5
6	8417	96.3	2596	14	US-10-307-019-6
7	8403	96.1	1618	12	US-10-182-243-46
8	8399	96.1	1610	14	US-10-307-019-4
9	7069	80.9	1351	14	US-10-307-019-1
10	4533	51.9	871	16	US-10-311-034-20
11	4529	51.8	871	14	US-10-307-019-7
12	2447.5	28.0	548	14	US-10-307-019-8
13	2415.5	27.6	548	14	US-10-307-019-9
14	2045	23.4	390	12	US-10-425-114-37530
15	1973	22.6	2380	12	US-10-333-314-18

ALIGNMENTS

RESULT 1

US-09-858-664A-2
; Sequence 2, Application US/09858664A
; Patent No. US20020072491A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-2

Query Match 100.0%; Score 8740; DB 9; Length 1665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGCCRLGCGCVSAHVSQGLTNHPSVWGWHPLGCGGGLHSSLPALPPPSMQVTI 60
Db 1 MGCCRLGCGCVSAHVSQGLTNHPSVWGWHPLGCGGGLHSSLPALPPPSMQVTI 60
Qy 61 EDVQAQTGTAQFEAIIIRGDPQPSVTWKDSVQLVDSRLSQQEGTTSYSLVRHASKD 120
Db 61 EDVQAQTGTAQFEAIIIRGDPQPSVTWKDSVQLVDSRLSQQEGTTSYSLVRHASKD 120
Qy 121 AGVYTCLAQNTGQVLCRAELLVLGDDNEPDSKSHRKLHSFYEVKEETGRGVFGVK 180
Db 121 AGVYTCLAQNTGQVLCRAELLVLGDDNEPDSKSHRKLHSFYEVKEETGRGVFGVK 180

QY 181 RVQKGNKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETKTLILLEL 240
DB 181 RVQKGNKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETKTLILLEL 240
QY 241 CSSELDRLRYKGVVTEAVKVIYIQQVLEGLYHSHGVHLHDIKPSNLMVHPAREDI 300
DB 241 CSSELDRLRYKGVVTEAVKVIYIQQVLEGLYHSHGVHLHDIKPSNLMVHPAREDI 300
QY 301 KICDFGAQNTTFAELQFQSGYGFVSPBIIQQNFVSEASDIWANGVISYLSLTCSSPF 360
DB 301 KICDFGAQNTTFAELQFQSGYGFVSPBIIQQNFVSEASDIWANGVISYLSLTCSSPF 360
QY 361 AGESDRATLNLVLEGRVSWSPAAHLSEDAKDFIKATLQAPQAPPSAAQCLSHPWELK 420
DB 361 AGESDRATLNLVLEGRVSWSPAAHLSEDAKDFIKATLQAPQAPPSAAQCLSHPWELK 420
QY 421 SMPAEBAHFINTKQLKELLARSQWRLMSYKSLVMSRIPBELLRGPPDPSLGVARHLK 480
DB 421 SMPAEBAHFINTKQLKELLARSQWRLMSYKSLVMSRIPBELLRGPPDPSLGVARHLK 480
QY 481 RDTGSSSSSSSSDNLAPARAKSLPPSPVTHSPHLLHPRGTFRLPSASLPEBAEASERST 540
DB 481 RDTGSSSSSSSSDNLAPARAKSLPPSPVTHSPHLLHPRGTFRLPSASLPEBAEASERST 540
QY 541 EAPAPPASPGAGPPAAQGCVPKHSVIRSLFYHQAGESPEHGALAPGRRRHPPARRHLK 600
DB 541 EAPAPPASPGAGPPAAQGCVPKHSVIRSLFYHQAGESPEHGALAPGRRRHPPARRHLK 600
QY 601 GGYIAGALPGLRBLPMLHRVLEBAAREEQATLLAKAPFETALRLPASGTHLAPGHS 660
DB 601 GGYIAGALPGLRBLPMLHRVLEBAAREEQATLLAKAPFETALRLPASGTHLAPGHS 660
QY 661 LEHDSPTTPSPSEACGEAQLRPSAGGAPTRDMGHQPGSKQLPSTGHPGTAQPERPS 720
DB 661 LEHDSPTTPSPSEACGEAQLRPSAGGAPTRDMGHQPGSKQLPSTGHPGTAQPERPS 720
QY 721 PDSFWGQAPAFCHPKQGSAPQECSPHPAVAPCPGSPFPFGCKEAPLVPSSPFLGQQA 780
DB 721 PDSFWGQAPAFCHPKQGSAPQECSPHPAVAPCPGSPFPFGCKEAPLVPSSPFLGQQA 780
QY 781 PPAPAKASPPLDSKMGDITSLGRPKPGPCSPGASQASSQVSSLRVSSQVGTPEP 840
DB 781 PPAPAKASPPLDSKMGDITSLGRPKPGPCSPGASQASSQVSSLRVSSQVGTPEP 840
QY 841 PSIDAEGWTCQAEADLSDSTFTLQPOBQVTKRKFSLGGGGVAGVAGTFAFGDAGM 900
DB 841 PSIDAEGWTCQAEADLSDSTFTLQPOBQVTKRKFSLGGGGVAGVAGTFAFGDAGM 900
QY 901 LGQPMWARIAMAVSQQEBAEAEQSEBQAEAESELPQVSAAPVEVGRAPTR 960
DB 901 LGQPMWARIAMAVSQQEBAEAEQSEBQAEAESELPQVSAAPVEVGRAPTR 960
QY 961 SSPEPTWEDIGQVSLVQIRDLSDGDAEADTSLDISVDPAVNLSDLYDKYLPPEFM 1020
DB 961 SSPEPTWEDIGQVSLVQIRDLSDGDAEADTSLDISVDPAVNLSDLYDKYLPPEFM 1020
QY 1021 IFRKVPKSAQPEPPSPMAEBELAEFPPTWPMFPGELGPHAGLEITESESDVALLAEAAV 1080
DB 1021 IFRKVPKSAQPEPPSPMAEBELAEFPPTWPMFPGELGPHAGLEITESESDVALLAEAAV 1080
QY 1081 GRKXKMSPPSRSLFHPGRHLPLDEPAELGLRERVKASVEHLSRLKGRPEGEKEGPPR 1140
DB 1081 GRKXKMSPPSRSLFHPGRHLPLDEPAELGLRERVKASVEHLSRLKGRPEGEKEGPPR 1140
QY 1141 KPGGLASFRSLGSKWDRAPTFLRELSDETIVLGGSVTLACQVSAQPAQATWSKDGAPL 1200
DB 1141 KPGGLASFRSLGSKWDRAPTFLRELSDETIVLGGSVTLACQVSAQPAQATWSKDGAPL 1200
QY 1201 ESSSRVLISATLKNFOLLITLVVAEDDLGVYTCVSNALGTITVTGVLKKAERPSSSPCP 1260
DB 1201 ESSSRVLISATLKNFOLLITLVVAEDDLGVYTCVSNALGTITVTGVLKKAERPSSSPCP 1260
QY 1261 DIGEVADGVLLVWKVPSYGPVTYIVQCSLEGGSWTTLASDIFDCCVLTSLKSRGGTYT 1320

DB 1261 DIGEVADGVLLVWKVPSYGPVTYIVQCSLEGGSWTTLASDIFDCCVLTSLKSRGGTYT 1320
QY 1321 FRTACVSKAGMCPYSPSEQVLLGSPSHLASEESQBSAQPLSTKTFAQTQIQGRF 1380
DB 1321 FRTACVSKAGMCPYSPSEQVLLGSPSHLASEESQBSAQPLSTKTFAQTQIQGRF 1380
QY 1381 SVVRQWEKASGRALAAKIIIPYHPKDKTAVLRVEYALKGLRHPHLAQHAAAYLSPRHVL 1440
DB 1381 SVVRQWEKASGRALAAKIIIPYHPKDKTAVLRVEYALKGLRHPHLAQHAAAYLSPRHVL 1440
QY 1441 ILELCSGPPELLCLABRASYSSEYKDYLMOWLSATQVYLNHNLHLDLRSNNMIITEYN 1500
DB 1441 ILELCSGPPELLCLABRASYSSEYKDYLMOWLSATQVYLNHNLHLDLRSNNMIITEYN 1500
QY 1501 LKWDVLDGNAQSLQSKVLPDKFYLETMAPELLEGQGAVPQTDIWAIGVTAFIMLSA 1560
DB 1501 LKWDVLDGNAQSLQSKVLPDKFYLETMAPELLEGQGAVPQTDIWAIGVTAFIMLSA 1560
QY 1561 EYVPSSEGARDLQGRGLVRLSCYAGLGGAVAFRLSTLCAQKGRPCASSCLOC PW 1620
DB 1561 EYVPSSEGARDLQGRGLVRLSCYAGLGGAVAFRLSTLCAQKGRPCASSCLOC PW 1620
QY 1621 LTEEGPACSRPAPVTFTPTARLRVFNREKRRALLYKRNHNLQVR 1665
DB 1621 LTEEGPACSRPAPVTFTPTARLRVFNREKRRALLYKRNHNLQVR 1665

RESULT 2
US-10-697-263-2
; Sequence 2, Application US/10697263
; Publication No. US20040063142A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-697-263-2

Query Match 100.0%; Score 8740; DB 12; Length 1665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCCRLGCGGCSVAHSVSGQLTNHPSVMVCGWHPGLCGWGGGLHSSLPALPSPSMQVTI 60
DB 1 MCCRLGCGGCSVAHSVSGQLTNHPSVMVCGWHPGLCGWGGGLHSSLPALPSPSMQVTI 60
QY 61 EDVQAGTGTGAQFAELIEGDPPQSVTWYKOSVQLVDSTRLSQQEGTYSYLVLRHVASKD 120
DB 61 EDVQAGTGTGAQFAELIEGDPPQSVTWYKOSVQLVDSTRLSQQEGTYSYLVLRHVASKD 120
QY 121 AGVYTCLAQNTGGQVLCXAEILLVIGGNEPDEKSHRRKLHSPYVEKEIGRGVFGVK 180
DB 121 AGVYTCLAQNTGGQVLCXAEILLVIGGNEPDEKSHRRKLHSPYVEKEIGRGVFGVK 180
QY 181 RVQKGNKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETKTLILLEL 240
DB 181 RVQKGNKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETKTLILLEL 240

QY 241 CSSELDRLVRKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKPSNLMVHPAREDI 300
Db 241 CSSELDRLVRKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKPSNLMVHPAREDI 300
QY 301 KICDFGFAQNTITPAELOFSQYSGSEFVSPETIQONPVSEASDIWAMGVISYLSLTCSPP 360
Db 301 KICDFGFAQNTITPAELOFSQYSGSEFVSPETIQONPVSEASDIWAMGVISYLSLTCSPP 360
QY 361 AGESDRATLLNVLEGRVSWSPMAHLSEDAKOFIKATLORAPQAPSPAAQCLSHWFLK 420
Db 361 AGESDRATLLNVLEGRVSWSPMAHLSEDAKOFIKATLORAPQAPSPAAQCLSHWFLK 420
QY 421 SMPAEAAHFINTKOLKFLLAARSWORSIMSYKSTILVWRSIPELLRGPPDPSPSLGVARHL 480
Db 421 SMPAEAAHFINTKOLKFLLAARSWORSIMSYKSTILVWRSIPELLRGPPDPSPSLGVARHL 480
QY 481 RDTGSS 540
Db 481 RDTGSS 540
QY 541 EAPAPPASPEGAGPPAQQGVPRHSVIRSLFYHQGESPEHGALAPGSRHHPARRRHLLK 600
Db 541 EAPAPPASPEGAGPPAQQGVPRHSVIRSLFYHQGESPEHGALAPGSRHHPARRRHLLK 600
QY 601 GGYTAGALPGLRPEIMHRVLEBAABEQATLLAKAPSPETALRLPASGTHLAPGHSHS 660
Db 601 GGYTAGALPGLRPEIMHRVLEBAABEQATLLAKAPSPETALRLPASGTHLAPGHSHS 660
QY 661 LEHDSPTSPSS 720
Db 661 LEHDSPTSPSS 720
QY 721 PDSWQGPAPCHPKQSAQOEGGSPHAPVAPCPGSPFPGSCKEAPLVPSPFLQPOA 780
Db 721 PDSWQGPAPCHPKQSAQOEGGSPHAPVAPCPGSPFPGSCKEAPLVPSPFLQPOA 780
QY 781 PPAKASPPDLDSXMGPGDLSLQRPKPGPCSPGSGASQSSQVSSSLRVGSSQVGTGEG 840
Db 781 PPAKASPPDLDSXMGPGDLSLQRPKPGPCSPGSGASQSSQVSSSLRVGSSQVGTGEG 840
QY 841 PSLDAEGTQBAEDLSSTPTLQRPQVQVTRKPSLGGRGYAGVAGYGTTFAGGDAGM 900
Db 841 PSLDAEGTQBAEDLSSTPTLQRPQVQVTRKPSLGGRGYAGVAGYGTTFAGGDAGM 900
QY 901 LGQPMWARIWAYSQGBEEQEEARASQSEBQBARABSPLOVYARVPVEVGRAPTR 960
Db 901 LGQPMWARIWAYSQGBEEQEEARASQSEBQBARABSPLOVYARVPVEVGRAPTR 960
QY 961 SSPFTPWEDIGQVSLVQIRDLSDGAEADTISLIDISEVDPAVNLSDLYDKYLPFFPM 1020
Db 961 SSPFTPWEDIGQVSLVQIRDLSDGAEADTISLIDISEVDPAVNLSDLYDKYLPFFPM 1020
QY 1021 IFRKVPKSAQEPSPMAEELAEFPPTWPMGELPHAGLEITRESEVDVALLAEAAV 1080
Db 1021 IFRKVPKSAQEPSPMAEELAEFPPTWPMGELPHAGLEITRESEVDVALLAEAAV 1080
QY 1081 GRKRWSSPSLSLPHFFGRHLPLDEPAELGLRVRKASVEHISRIILKGRPEGLEKGGPR 1140
Db 1081 GRKRWSSPSLSLPHFFGRHLPLDEPAELGLRVRKASVEHISRIILKGRPEGLEKGGPR 1140
QY 1141 KKPGLASFLSGLKSWDRAPTFLELSDTETVVLQSVTLACQVSAQAPAAQATWSKDGAPL 1200
Db 1141 KKPGLASFLSGLKSWDRAPTFLELSDTETVVLQSVTLACQVSAQAPAAQATWSKDGAPL 1200
QY 1201 ESSSRVLISATLKNFOLLTLVVAEDLGVYTCVSNALGTVTGVLKRAERPSSSPCP 1260
Db 1201 ESSSRVLISATLKNFOLLTLVVAEDLGVYTCVSNALGTVTGVLKRAERPSSSPCP 1260
QY 1261 DIGEVYADGVLLVWKPVSYPVTYIYQCSLEGGSWTTLASDIFDCCYLTSLKSRGGT 1320
Db 1261 DIGEVYADGVLLVWKPVSYPVTYIYQCSLEGGSWTTLASDIFDCCYLTSLKSRGGT 1320

QY 1321 FRTACVSKAGMGYSPPSEBOVLLGGPSHLASEBESQGRSAQPLPSTTKTAFQTOIGRGP 1380
Db 1321 FRTACVSKAGMGYSPPSEBOVLLGGPSHLASEBESQGRSAQPLPSTTKTAFQTOIGRGP 1380
QY 1381 SVVROCKEASGRALAAKIIPIYHPKDKTAVLREYEAALKGLRHPHLAQLHAAVLSPPHLVL 1440
Db 1381 SVVROCKEASGRALAAKIIPIYHPKDKTAVLREYEAALKGLRHPHLAQLHAAVLSPPHLVL 1440
QY 1441 ILELCSPPELLPCLASRASYSSEVYKWLQMLSATQYILHNOHILHLDRSENMIITEYN 1500
Db 1441 ILELCSPPELLPCLASRASYSSEVYKWLQMLSATQYILHNOHILHLDRSENMIITEYN 1500
QY 1501 LLKVVLDLGNAAQSLSQEKVLPDSKFKDYLETMAPELLEGGQAVPQTDIWAIGVTAFTMLSA 1560
Db 1501 LLKVVLDLGNAAQSLSQEKVLPDSKFKDYLETMAPELLEGGQAVPQTDIWAIGVTAFTMLSA 1560
QY 1561 EYPVSSSGARDLQRLKGLVRLSRVAGLSGAVAFIRSTICAQWGRPCASSCLOCPW 1620
Db 1561 EYPVSSSGARDLQRLKGLVRLSRVAGLSGAVAFIRSTICAQWGRPCASSCLOCPW 1620
QY 1621 LITEGPAACSPAPVPTPTARLRFVVRNREKRALLYKRNHLAQVR 1665
Db 1621 LITEGPAACSPAPVPTPTARLRFVVRNREKRALLYKRNHLAQVR 1665

RESULT 3

US-10-415-011-22
; Sequence 22, Application US/10415011
; Publication No. US20040053394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah K.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuming
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LAL, Preeti G.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KEAN, Farrah A.
; APPLICANT: KEAN, Farrah A.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; PRIOR FILING DATE: 2003-04-18
; PRIOR FILING DATE: 2001-10-20
; PRIOR FILING DATE: 2001-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 2000-11-03
; PRIOR FILING DATE: 2000-11-09
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/242,410
; PRIOR APPLICATION NUMBER: US 60/244,068
; PRIOR APPLICATION NUMBER: US 60/245,708
; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR APPLICATION NUMBER: US 60/249,565

;; PRIOR FILING DATE: 2001-02-15
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 2630
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-077-130-2

Query Match 96.4%; Score 8423; DB 13; Length 2630;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 PPSQVITIEDVQAGTGAQFEALIEGDPOPSVTWYKDSVOLVDSTLSQQQEGTYSLV 112
Db 1018 PPSQVITIEDVQAGTGAQFEALIEGDPOPSVTWYKDSVOLVDSTLSQQQEGTYSLV 1077

QY 113 LRHVASKDAGVYTCLAQNTGQVCKAEILLVLDGNEPDSEKSHRKLHSFVYKKEIG 172
Db 1078 LRHVASKDAGVYTCLAQNTGQVCKAEILLVLDGNEPDSEKSHRKLHSFVYKKEIG 1137

QY 173 RGVFGFVKRVQHKNGKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETR 232
Db 1138 RGVFGFVKRVQHKNGKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETR 1197

QY 233 TLIIILELCSSEELDLRYRGVVTAEARKYIIOQLVEGLHYLHSHGVHLHDIKPSNIM 292
Db 1198 TLIIILELCSSEELDLRYRGVVTAEARKYIIOQLVEGLHYLHSHGVHLHDIKPSNIM 1257

QY 293 VHPAREDIKICDFGFAQNTIPEALQFSGPSPVPEIIQQNPVSEASDIWANGVISYL 352
Db 1258 VHPAREDIKICDFGFAQNTIPEALQFSGPSPVPEIIQQNPVSEASDIWANGVISYL 1317

QY 353 SLTCSPPAGSDRATLNLVLEGRVSWSSPMAHLSEDAKFIKATLQAPQAPPSAAQC 412
Db 1318 SLTCSPPAGSDRATLNLVLEGRVSWSSPMAHLSEDAKFIKATLQAPQAPPSAAQC 1377

QY 413 LSHPFLKSMABEAHFNTKQLFLARSQWLSMSYKILVMSIPELLRPPDPS 472
Db 1378 LSHPFLKSMABEAHFNTKQLFLARSQWLSMSYKILVMSIPELLRPPDPS 1437

QY 473 LGVARHLCRDTGGSSSSSSSDNELAPPARAKSLPPSPVTHSLPHRPFRLRPSASLPPE 532
Db 1438 LGVARHLCRDTGGSSSSSSSDNELAPPARAKSLPPSPVTHSLPHRPFRLRPSASLPPE 1497

QY 533 AEASRSTEAAPAPSPAGGAPPAAGCVPRHSVIRSLFYHQAGSEPHGALAPGSRHP 592
Db 1498 AEASRSTEAAPAPSPAGGAPPAAGCVPRHSVIRSLFYHQAGSEPHGALAPGSRHP 1557

QY 593 ARRRHLKGGYIAGALPGLREPLMEHRVLEBEAAREEQATLLAKAPFETALRLPASGTH 652
Db 1558 ARRRHLKGGYIAGALPGLREPLMEHRVLEBEAAREEQATLLAKAPFETALRLPASGTH 1617

QY 653 LAPGSHSLHSDSPTRPSSEACGEAQLPSAGGAPIRDMGHPOGSKOLPSTGGHPG 712
Db 1618 LAPGSHSLHSDSPTRPSSEACGEAQLPSAGGAPIRDMGHPOGSKOLPSTGGHPG 1677

QY 713 TQAPERPDSPWGPAPFCHPKQGSAPQEGCSHPAVAPCPGSPFGSCKEAPLPVPS 772
Db 1678 TQAPERPDSPWGPAPFCHPKQGSAPQEGCSHPAVAPCPGSPFGSCKEAPLPVPS 1737

QY 773 PFLGQFQAPPAPAKASPLDLSKMGEDISLPGRPKPGPCSPFGSASQASSQVSLRVGS 832
Db 1738 PFLGQFQAPPAPAKASPLDLSKMGEDISLPGRPKPGPCSPFGSASQASSQVSLRVGS 1797

QY 833 SOVGTEPGSLDAEGWTOEADLSSTPTLQRPQVQVTRKFSLGGRGYAGVAGYGTFA 892
Db 1798 SOVGTEPGSLDAEGWTOEADLSSTPTLQRPQVQVTRKFSLGGRGYAGVAGYGTFA 1857

QY 893 FGGDAGGMLGQGMWARIWAYSQSEEEQEEARAESQSEEQEAPARASPLPQVSARVP 952
Db 1858 FGGDAGGMLGQGMWARIWAYSQSEEEQEEARAESQSEEQEAPARASPLPQVSARVP 1917

QY 953 EVGRAPTRSSPEPTWEDIGQVSLVQIIRDLSDGAFAADTISLDISEVDPAYNLSDLYDI 1012
Db 1918 EVGRAPTRSSPEPTWEDIGQVSLVQIIRDLSDGAFAADTISLDISEVDPAYNLSDLYDI 1977

QY 1013 KYLPEEFMIERKVPKSAQPEPPSMAEELAEFPPTWPMCELGPHAGLEITSESDVD 1072
Db 1978 KYLPEEFMIERKVPKSAQPEPPSMAEELAEFPPTWPMCELGPHAGLEITSESDVD 2037

QY 1073 ALLAAAVGRKRWSSPSRSLFHPFGRHLPLDPAELGLRERVKASVEHISRIILKGRPEG 1132
Db 2038 ALLAAAVGRKRWSSPSRSLFHPFGRHLPLDPAELGLRERVKASVEHISRIILKGRPEG 2097

QY 1133 LEKEGPPRRKPGGLASFRLSGLKSWDRAPTEFLRELSDETVLQSVVTLACQVSAQPAQAAT 1192
Db 2098 LEKEGPPRRKPGGLASFRLSGLKSWDRAPTEFLRELSDETVLQSVVTLACQVSAQPAQAAT 2157

QY 1193 WSKDGAPLESSRVLISATLKNFQLLTILVVAEDLGYITCSVSNALGTVTGTVLRKAE 1252
Db 2158 WSKDGAPLESSRVLISATLKNFQLLTILVVAEDLGYITCSVSNALGTVTGTVLRKAE 2217

QY 1253 RPSSSPCPDIGEVADGVLVWKPVSYPVTYIQCSEGGSWTTLASDIFDCCVLTSK 1312
Db 2218 RPSSSPCPDIGEVADGVLVWKPVSYPVTYIQCSEGGSWTTLASDIFDCCVLTSK 2277

QY 1313 LSRGGTYTPTACVSKAGMPYSSPSEQVLLGPGSHLASEEESQGRSAQPLSTKTFAFQ 1372
Db 2278 LSRGGTYTPTACVSKAGMPYSSPSEQVLLGPGSHLASEEESQGRSAQPLSTKTFAFQ 2337

QY 1373 TQIQGRFVSVQCKEASGRALAAKIIPYHPKDKTAVLREYEAALKGLRPHLAOLHAAY 1432
Db 2338 TQIQGRFVSVQCKEASGRALAAKIIPYHPKDKTAVLREYEAALKGLRPHLAOLHAAY 2397

QY 1433 LSPRHVLVLELCSGPELLPCLAEASYSSEVKDYLMQMLSATQVYLNQHLHLDRSE 1492
Db 2398 LSPRHVLVLELCSGPELLPCLAEASYSSEVKDYLMQMLSATQVYLNQHLHLDRSE 2457

QY 1493 NMITEYNLLKVVLDGNAOSLSQEKVLPDSKFKDYLETWAPELLEGGQAVPQTDIWAIGV 1552
Db 2458 NMITEYNLLKVVLDGNAOSLSQEKVLPDSKFKDYLETWAPELLEGGQAVPQTDIWAIGV 2517

QY 1553 TAFIMLSAEVPSVSEGGARDLQRLKGLVRLSRVAGLGGAVAFURSTLCAQPMGRPCA 1612
Db 2518 TAFIMLSAEVPSVSEGGARDLQRLKGLVRLSRVAGLGGAVAFURSTLCAQPMGRPCA 2577

QY 1613 SSCLOCPWLTEEGPACSRPAPVTFPTARLVFVRNREKRALLYKHNLAQVR 1665
Db 2578 SSCLOCPWLTEEGPACSRPAPVTFPTARLVFVRNREKRALLYKHNLAQVR 2630

RESULT 5
US-10-077-130-5
; Sequence 5, Application US/100771130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MEI2001-0471RCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 7968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-130-5

Query Match 96.4%; Score 8423; DB 13; Length 7968;
Best Local Similarity 99.9%; Pred. No. 0;

QY 173 RGVFGVVRVQKGNKILCAAKFIPLSRTRAQYRERDILAAALSHPLVLTGLDQFETRK 232
 Db 1104 RGVFGVVRVQKGNKILCAAKFIPLSRTRAQYRERDILAAALSHPLVLTGLDQFETRK 1163
 QY 233 TLIIILELCSSEILLDRILRGVGVTEAEVKYIIOQLVEGLHYLHSHGVHLHDIKPSNIML 292
 Db 1164 TLIIILELCSSEILLDRILRGVGVTEAEVKYIIOQLVEGLHYLHSHGVHLHDIKPSNIML 1223
 QY 293 VHPAREDIKICDFGAQNIITAEIQLFQSGPESFVPSPEIIOQNPVSBASDIWANGVISYL 352
 Db 1224 VHPAREDIKICDFGAQNIITAEIQLFQSGPESFVPSPEIIOQNPVSBASDIWANGVISYL 1283
 QY 353 SLTCSPPFAGESDRATLNLVLEGRVSMSPMAHLSEDAKFIKATIQAPQAPPSAAQC 412
 Db 1284 SLTCSPPFAGESDRATLNLVLEGRVSMSPMAHLSEDAKFIKATIQAPQAPPSAAQC 1343
 QY 413 LSHPWFLKSMFAEEHFTNTQLKFLARSQWLSMSYKILVMSRSTPELIRGPPDPS 472
 Db 1344 LSHPWFLKSMFAEEHFTNTQLKFLARSQWLSMSYKILVMSRSTPELIRGPPDPS 1403
 QY 473 LGVARHLCRDTCGSSSSSSSDNELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPPEE 532
 Db 1404 LGVARHLCRDTCGSSSSSSSDNELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPPEE 1463
 QY 533 AEASERSTAPASPAGAPPAAGCVPVHSVIRSLFYHOAGHSPHGGALAGSRHP 592
 Db 1464 AEASERSTAPASPAGAPPAAGCVPVHSVIRSLFYHOAGHSPHGGALAGSRHP 1523
 QY 593 ARRRHLKGGYTAGALPGLRPLRPLMRHVRLEBAAREQATLLAKAPSPETALRLPASGTH 652
 Db 1524 ARRRHLKGGYTAGALPGLRPLRPLMRHVRLEBAAREQATLLAKAPSPETALRLPASGTH 1593
 QY 653 LAPGSHSLHSDPSTPRPSSAOCQALRPSAPSGGAPIRDMGHPQSGKOLPSTGGHPG 712
 Db 1594 LAPGSHSLHSDPSTPRPSSAOCQALRPSAPSGGAPIRDMGHPQSGKOLPSTGGHPG 1643
 QY 713 TAQPERPSPDPSWGPAPFCHPKQSAQEGCSHPAVAPCPGSPGSCKEAPLPVSS 772
 Db 1644 TAQPERPSPDPSWGPAPFCHPKQSAQEGCSHPAVAPCPGSPGSCKEAPLPVSS 1703
 QY 773 PFLGQFQAPPAPAKASPLDQKMGPGDILSLGRPKPGPCSPGSAQSSQVSSLRVGS 832
 Db 1704 PFLGQFQAPPAPAKASPLDQKMGPGDILSLGRPKPGPCSPGSAQSSQVSSLRVGS 1763
 QY 833 SQVGTPEGSLDAEGWTQAEPLDSPTLQRPQEQVTRKFSLLGGRGQYAGVAGYGA 892
 Db 1764 SQVGTPEGSLDAEGWTQAEPLDSPTLQRPQEQVTRKFSLLGGRGQYAGVAGYGA 1823
 QY 893 FGDAGGMLGQGPMMARIWAVYSQSEEEQEBARAESQSEEQEAPAESPLPQVSARVPV 952
 Db 1824 FGDAGGMLGQGPMMARIWAVYSQSEEEQEBARAESQSEEQEAPAESPLPQVSARVPV 1883
 QY 953 EVGRAPTRSPPTWEDIGQVSLQVIRDLSDGAEADTISLDISEVDPAYINLSLDYDI 1012
 Db 1884 EVGRAPTRSPPTWEDIGQVSLQVIRDLSDGAEADTISLDISEVDPAYINLSLDYDI 1943
 QY 1013 KYLPPEFMIKRVKPSAQPEPPSPMAEEELAEFPPTWPCGELGPHAGLEITESESDVD 1072
 Db 1944 KYLPPEFMIKRVKPSAQPEPPSPMAEEELAEFPPTWPCGELGPHAGLEITESESDVD 2003
 QY 1073 ALLAAVAVKRWKSPSSSLFHPGRHPLDEPAELGLRERVKASVEHISRLKGRPEG 1132
 Db 2004 ALLAAVAVKRWKSPSSSLFHPGRHPLDEPAELGLRERVKASVEHISRLKGRPEG 2063
 QY 1133 LEKEGPRKPKGLASPRISGLKSWDRAPTFRLSDSETVVLGQSVTLACOVSAQAQAT 1192
 Db 2064 LEKEGPRKPKGLASPRISGLKSWDRAPTFRLSDSETVVLGQSVTLACOVSAQAQAT 2123
 QY 1193 WSKDGAFLSSSRVLISATLKNFQLLTILVVAEDLGVTCTSVSNALGVTTTGVLRKAE 1252
 Db 2124 WSKDGAFLSSSRVLISATLKNFQLLTILVVAEDLGVTCTSVSNALGVTTTGVLRKAE 2183

QY 1253 RPSSSPCPDPIGBVYADGVLLVWKPVSYPVTVIQCSELEGSSWTTLASDIFDCCYLTSK 1312
 Db 2184 RPSSSPCPDPIGBVYADGVLLVWKPVSYPVTVIQCSELEGSSWTTLASDIFDCCYLTSK 2243
 QY 1313 LSRGGTYTERTACVSKAGMPYSSSEQVLLCGPSHLASEEESQGRSAQPLSTKTFAFQ 1372
 Db 2244 LSRGGTYTERTACVSKAGMPYSSSEQVLLCGPSHLASEEESQGRSAQPLSTKTFAFQ 2303
 QY 1373 TQIQGRFSPVVRQCKEASGRALAAKIIPYHPKDKTAVLRREYALKGLRHPHLAQHAA 1432
 Db 2304 TQIQGRFSPVVRQCKEASGRALAAKIIPYHPKDKTAVLRREYALKGLRHPHLAQHAA 2363
 QY 1433 LSPRHLVLTLELCSGPPELLPCLAEASYSSEVKDYLMQMLSATOVLYHNOHILHLDLRSE 1492
 Db 2364 LSPRHLVLTLELCSGPPELLPCLAEASYSSEVKDYLMQMLSATOVLYHNOHILHLDLRSE 2423
 QY 1493 NMIITEYNLLKVVLDLGNALQSQKYLPSDKFKDYLETNMAPELLBEGQAVPOTDIWAIGV 1552
 Db 2424 NMIITEYNLLKVVLDLGNALQSQKYLPSDKFKDYLETNMAPELLBEGQAVPOTDIWAIGV 2483
 QY 1553 TAPIMLSAEPYSSSEGARDLQRLKGLVRLSRCVAGLSGGAVAFRLSTLCAQPMWRPCCA 1612
 Db 2484 TAPIMLSAEPYSSSEGARDLQRLKGLVRLSRCVAGLSGGAVAFRLSTLCAQPMWRPCCA 2543
 QY 1613 SSCLOCPMLTERGPACSRPAPVTFPTARLVRVVRNREKRRALLYKEHNLAQVR 1665
 Db 2544 SSCLOCPMLTERGPACSRPAPVTFPTARLVRVVRNREKRRALLYKEHNLAQVR 2596

RESULT 7
 US-10-182-243-46
 ; Sequence 46, Application US/10182243
 ; Publication No. US20040048310A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PLOWMAN, GREGORY D.
 ; APPLICANT: WHYTE, DAVID
 ; APPLICANT: MANNING, GERARD
 ; APPLICANT: SUDARSANAM, SUCHA
 ; APPLICANT: MARTINEZ, RICARDO
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
 ; TITLE OF INVENTION: ENZYMES
 ; FILE REFERENCE: 038602/1366
 ; CURRENT APPLICATION NUMBER: US/10/182,243
 ; CURRENT FILING DATE: 2003-07-07
 ; PRIOR APPLICATION NUMBER: PCT/US01/02337
 ; PRIOR FILING DATE: 2001-01-25
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 46
 ; LENGTH: 1618
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-182-243-46

Query Match 96.1%; Score 8403; DB 12; Length 1618;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 161; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 54 PSMQVTTEDVQAGTGGTAQFEAIEGDPQSPVWYKDSVLVDSTRLSQOQEGTYSVLV 113
 Db 1 PSMQVTTEDVQAGTGGTAQFEAIEGDPQSPVWYKDSVLVDSTRLSQOQEGTYSVLV 60
 QY 114 RHVASKDAGVYTCIAQNTGGVLCCKAELLVLG-----GDNEPDSEKSHRRKLHSFYEV 167
 Db 61 RHVASKDAGVYTCIAQNTGGVLCCKAELLVLGASHSLGDNEDSEKSHRRKLHSFYEV 120
 QY 168 KEEIGRGVFGVVRVQKGNKILCAAKFIPLSRTRAQYRERDILAAALSHPLVLTGLDQ 227
 Db 121 KEEIGRGVFGVVRVQKGNKILCAAKFIPLSRTRAQYRERDILAAALSHPLVLTGLDQ 180
 QY 228 FETKTLILILELCSSEILLDRILRGVGVTEAEVKYIIOQLVEGLHYLHSHGVHLHDIKP 287
 Db 161 FETKTLILILELCSSEILLDRILRGVGVTEAEVKYIIOQLVEGLHYLHSHGVHLHDIKP 240

Db	1441	ITYENLLKVVDLGNAQSLSEQKVLPSDFDKFDYLETMAPELLSQGQAVPQTDIWAIGTAF	1500
Qy	1556	IMLSAIPVSSREGARDLQGLRAKGLVRLSRCVAGLSGGGAVAFRLSTLCAPQWGRPCASS	1615
Qy	1501	IMLSAIPVSSREGARDLQGLRAKGLVRLSRCVAGLSGGGAVAFRLSTLCAPQWGRPCASS	1560
Qy	1616	LQCPWLITERGPACSRPAPVTFTARLVFVRNREKERALLYKRNLAQVR	1665
Db	1561	LQCPWLITERGPACSRPAPVTFTARLVFVRNREKERALLYKRNLAQVR	1610

	Query Match	80.9%	Score 7069	DB 14	Length 1351	
	Best Local Similarity 99.9%	Pred. No. 0				
	Matches 1346	Conservative 1	Mismatches 1	Indels 0	Gaps 0	
QY	316	LQFSQYGSPEFVSPETIIQQNPVSEASDIWAGVISYLSLTCSPPFAGESDRATLLNVLEG	375			
DB	2	VQFSQYGSPEFVSPETIIQQNPVSEASDIWAGVISYLSLTCSPPFAGESDRATLLNVLEG	61			
QY	376	RVSWSSPMAHLSEDKADFATLQAPQAPPSAAOCLSHPWFLKSNPAEAEHFINTKQL	435			
DB	62	RVSWSSPMAHLSEDKADFATLQAPQAPPSAAOCLSHPWFLKSNPAEAEHFINTKQL	121			
QY	436	KFILLARSRWQSRSLMSYKSIILWMSIPELLRGPPDSPSLGVARHLCRDTGSSSSSSSSDN	495			
DB	122	KFILLARSRWQSRSLMSYKSIILWMSIPELLRGPPDSPSLGVARHLCRDTGSSSSSSSSDN	181			
QY	496	ELAPPAPAKSIPSPSVTHSLPILLHPRGFLRPSASIPPEAEASERSTERAPAPASPEGAGPP	555			
DB	182	ELAPPAPAKSLPPSVTHSLPILLHPRGFLRPSASIPPEAEASERSTERAPAPASPEGAGPP	241			
QY	556	AAQCGVPRHSVIRSLFYHQAGESPEHGALAPGRRHPARRRHLLKGYIAGALPCLREPL	615			
DB	242	AAQCGVPRHSVIRSLFYHQAGESPEHGALAPGRRHPARRRHLLKGYIAGALPGLREPL	301			
QY	616	MEHVLVEEAREEQATLLAKASFEFATALRPSAGTHLAPGHSLSLHSDSPSTPRPSSEA	675			
DB	302	MEHVLVEEAREEQATLLAKASFEFATALRPSAGTHLAPGHSLSLHSDSPSTPRPSSEA	361			
QY	676	CGEAQRLPSPAGGAPITRDMGHGPGSKQLPSTGTHGPTAQPERSPSPSPWGQAPAPFCHPK	735			
DB	362	CGEAQRLPSPAGGAPITRDMGHGPGSKQLPSTGTHGPTAQPERSPSPSPWGQAPAPFCHPK	421			
QY	736	QGSAPQGCSCSHPAVACPGRSPPGSCKEAPLVPSSPFLGQOPAPAPAKASPLDLSKM	795			
DB	422	QGSAPQGCSCSHPAVACPGRSPPGSCKEAPLVPSSPFLGQOPAPAPAKASPLDLSKM	481			
QY	796	GPFGISLIPGRPKPFCSPGSPGASQASSQSVSSLRVGSSQVGTGEPGSLDAGWTQAEADL	855			

482 GFGDISLPGKPGKPGSCGSAQSSQVSSLRVSSQVGTGPGPSLDAEGWTQAEADL 541
QY 856 SDSTPTLQRPQEQVWTKRKSLSGRGYAGVAGTGFAGDAGGMLGQPMWARIAMAYS 915
Db 542 SDSTPTLQRPQEQATWTKRKSLSGRGYAGVAGTGFAGDAGGMLGQPMWARIAMAYS 601
QY 916 QSEEEQEEARASQSEEEQEEARASPLQVSGARPVEVGRAPTRSSPEPTWEDIGQVS 975
Db 602 QSEEEQEEARASQSEEEQEEARASPLQVSGARPVEVGRAPTRSSPEPTWEDIGQVS 661
QY 976 LVQIRLSDGDAEAADTISLDSVDPAYNLNLDYDKYLPPEFMIFRKVPKSAQPEPPS 1035
Db 662 LVQIRLSDGDAEAADTISLDSVDPAYNLNLDYDKYLPPEFMIFRKVPKSAQPEPPS 721
QY 1036 PMAEEELAEPEPTWMPGELGPHAGLEITESEDDVALLAAVGRKRWSSPSRSLFH 1095
Db 722 PMAEEELAEPEPTWMPGELGPHAGLEITESEDDVALLAAVGRKRWSSPSRSLFH 781
QY 1096 PGRHPLDPAELGIRERVKASVEHISRLKGRPEGLEKEGPPRKKGLASFRLSGLKS 1155
Db 782 PGRHPLDPAELGIRERVKASVEHISRLKGRPEGLEKEGPPRKKGLASFRLSGLKS 841
QY 1156 WDRAPFLRLSDTETVVLGQSVTLAQVSAQAAQATWSKCAPLESSRVLISATLKNF 1215
Db 842 WDRAPFLRLSDTETVVLGQSVTLAQVSAQAAQATWSKCAPLESSRVLISATLKNF 901
QY 1216 QLLTILVVAEDLGVTTCVSNALGVTGVLKRAERPSSPCPDIGEVADGVLLVWK 1275
Db 902 QLLTILVVAEDLGVTTCVSNALGVTGVLKRAERPSSPCPDIGEVADGVLLVWK 961
QY 1276 PVSYSQPVTVIVQCSLEGSSWTLASDIFDCCYLTSKLSRGTYTFRACVSKAGMGFYS 1335
Db 962 PVSYSQPVTVIVQCSLEGSSWTLASDIFDCCYLTSKLSRGTYTFRACVSKAGMGFYS 1021
QY 1336 SPSEQVLGSPSHLASEESQGRSAQPLSTKTFATQIQGRSVVVRQWCKASGRAL 1395
Db 1022 SPSEQVLGSPSHLASEESQGRSAQPLSTKTFATQIQGRSVVVRQWCKASGRAL 1081
QY 1396 AAKIIPYHPKDKTAVLREYALKGLRPHLAQLHAAYLSRHLVLILELCSGPELLPCLA 1455
Db 1082 AAKIIPYHPKDKTAVLREYALKGLRPHLAQLHAAYLSRHLVLILELCSGPELLPCLA 1141
QY 1456 ERASYSESEVKDYLWMLSATQVYLNHNOHILHDLSENMIITEYNLLKVVDLGNAQSLSQ 1515
Db 1142 ERASYSESEVKDYLWMLSATQVYLNHNOHILHDLSENMIITEYNLLKVVDLGNAQSLSQ 1201
QY 1516 EKVLPDKFKDYLETWAPLLEGGQAVPQTDIWAIGVTAFIMLSAEYFVSSEGARDLQRG 1575
Db 1202 EKVLPDKFKDYLETWAPLLEGGQAVPQTDIWAIGVTAFIMLSAEYFVSSEGARDLQRG 1261
QY 1576 LRKGLVRLSCYAGLSGGAVALRSTLCAQPNRCPACSSCLOCPLTEEGPACSRPAPVT 1635
Db 1262 LRKGLVRLSCYAGLSGGAVALRSTLCAQPNRCPACSSCLOCPLTEEGPACSRPAPVT 1321
QY 1636 FPTARLRFVVRNREKRALLYKHNLAQVR 1665
Db 1322 FPTARLRFVVRNREKRALLYKHNLAQVR 1351

RESULT 10
US-10-311-034-20
; Sequence 20, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.

APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: CHAMLA, Narinder K.
APPLICANT: YAO, Monique G.
APPLICANT: LU, Dyung Aina M.
APPLICANT: GREENWALD, Sara R.
APPLICANT: RAMKUMAR, Jayalakmi
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: KEARNEY, Liam
APPLICANT: BURFORD, Neil
APPLICANT: NGUYEN, Damiel B.
APPLICANT: TANG, Y. Iom
APPLICANT: BAUGHN, Mariah R.
APPLICANT: HE, Ann
APPLICANT: THORNTON, Michael
APPLICANT: HAFALIA, April
APPLICANT: ARVIZU, Chandra S.
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: LO, Terence P.
APPLICANT: KHAH, Farrah A.
APPLICANT: RECIPON, Shirley A.
APPLICANT: AZIMZAI, Yalda
APPLICANT: POLICKY, Jennifer L.
APPLICANT: DING, Li
APPLICANT: GREYHER, Megan
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: BATRA, Sajeew
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: HUMAN KINASES
FILE REFERENCE: PI-0125 PCT
CURRENT APPLICATION NUMBER: US/10/311,034
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
60/228,056
PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-07-15
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PERL Program
SEQ ID NO 20
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040023242A1 3016969CD1
US-10-311-034-20

Query Match 51.9%; Score 4533; DB 16; Length 871;
Best Local Similarity 100.0%; Pred. No. 1.9e-218;
Matches 871; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 MGPGDISLPGKPGKPGSCGSAQSSQVSSLRVSSQVGTGPGPSLDAEGWTQAEAD 854
Db 1 MGPGDISLPGKPGKPGSCGSAQSSQVSSLRVSSQVGTGPGPSLDAEGWTQAEAD 60
QY 855 LSDSTPTLQRPQEQVWTKRKSLSGRGYAGVAGTGFAGDAGGMLGQPMWARIAMAYS 914
Db 61 LSDSTPTLQRPQEQVWTKRKSLSGRGYAGVAGTGFAGDAGGMLGQPMWARIAMAYS 120
QY 915 SOSSEEEQEEARASQSEEEQEEARASPLQVSGARPVEVGRAPTRSSPEPTWEDIGQV 974
Db 121 SOSSEEEQEEARASQSEEEQEEARASPLQVSGARPVEVGRAPTRSSPEPTWEDIGQV 180
QY 975 SLVQIRLSDGDAEAADTISLDSVDPAYNLNLDYDKYLPPEFMIFRKVPKSAQPEPP 1034
Db 181 SLVQIRLSDGDAEAADTISLDSVDPAYNLNLDYDKYLPPEFMIFRKVPKSAQPEPP 240
QY 1035 SPMAEEELAEPEPTWMPGELGPHAGLEITESEDDVALLAAVGRKRWSSPSRSLF 1094
Db 241 SPMAEEELAEPEPTWMPGELGPHAGLEITESEDDVALLAAVGRKRWSSPSRSLF 300
QY 1095 HPPGSHLPLDPAELGIRERVKASVEHISRLKGRPEGLEKEGPPRKKGLASFRLSGLK 1154

Db 301 HFGRLPLDEPAELGLRERVKASVEHISRLKGRPEGLEKSGPPKPKGLASFLSLGLK 360
QY 1155 SWDRAPTFRLBELSDETVLQGSVTLACQVSAQPAQAATWSKOGAPLESRRVLSATLKN 1214
Db 361 SWDRAPTFRLBELSDETVLQGSVTLACQVSAQPAQAATWSKOGAPLESRRVLSATLKN 420
QY 1215 FQLLTILVVAEDLGVYTCVSNALGTVTGVLKAEPPSSPCPDIGEVYADGVLLVM 1274
Db 421 FQLLTILVVAEDLGVYTCVSNALGTVTGVLKAEPPSSPCPDIGEVYADGVLLVM 480
QY 1275 KPVSYPVTYIIVQCSLGGSWTTLASDIFDCCYLTSKLGRGTYTFTACVSKAGMGPY 1334
Db 481 KPVSYPVTYIIVQCSLGGSWTTLASDIFDCCYLTSKLGRGTYTFTACVSKAGMGPY 540
QY 1335 SSPSEQVLLGGPSHLASEEESQGRSAQPLPSTKTFQIQRGRFVSVRQCEKASGRA 1394
Db 541 SSPSEQVLLGGPSHLASEEESQGRSAQPLPSTKTFQIQRGRFVSVRQCEKASGRA 600
QY 1395 LAAKIIPYHPKDKTAVLREYALKGRHPHLAQLHAAYLSRHLVLIILELCSGPELLPCL 1454
Db 601 LAAKIIPYHPKDKTAVLREYALKGRHPHLAQLHAAYLSRHLVLIILELCSGPELLPCL 660
QY 1455 AERASYSSEVKDYLMQMLSATQYLHNOHILHDLRSENMIITEYNLLKVVLDLGNASLS 1514
Db 661 AERASYSSEVKDYLMQMLSATQYLHNOHILHDLRSENMIITEYNLLKVVLDLGNASLS 720
QY 1515 QEKVLPDSKDKYLETMAPELLEGQAVPQTDIWAIGVTAFLMIAEYFVSSEKARDLQR 1574
Db 721 QEKVLPDSKDKYLETMAPELLEGQAVPQTDIWAIGVTAFLMIAEYFVSSEKARDLQR 780
QY 1575 GLRKGVLRLSRVYAGLSGGAFLRSTLCAQPWGRPCASSCLOCPWLTTEGPACSRPAPV 1634
Db 781 GLRKGVLRLSRVYAGLSGGAFLRSTLCAQPWGRPCASSCLOCPWLTTEGPACSRPAPV 840
QY 1635 TPTARLVRVFNREKRALLYKRNHLAQR 1665
Db 841 TPTARLVRVFNREKRALLYKRNHLAQR 871

RESULT 11

US-10-307-019-7

; Sequence 7, Application US/10307019
; Publication No. US20030108533A1
; GENERAL INFORMATION:
; APPLICANT: Zeng, Wenlin
; APPLICANT: Stanton, Lawrence
; APPLICANT: SCIOS, INC.
; TITLE OF INVENTION: MEDIATORS OF SIGNAL TRANSDUCTION
; FILE REFERENCE: SCIOS.021DV1
; CURRENT APPLICATION NUMBER: US/10/307,019
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/548,473
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/129,552
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-019-7

Query Match 51.8%; Score 4529; DB 14; Length 871;
Best Local Similarity 99.9%; Pred. No. 2.9e-218;
Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 795 MGFGDLSLGRPKPGPCSPGASQSSQVSLRVGSSQVGTETPGPSLDAGWTQAEAD 854
Db 1 MGFGDLSLGRPKPGPCSPGASQSSQVSLRVGSSQVGTETPGPSLDAGWTQAEAD 60
QY 855 LSDSTFTLQRPQVTRKFSLGGRGYAGVAGYGTAFPGDGAGMLQGGPWWARIAWAV 914

Db 61 LSDSTFTLQRPQVTRKFSLGGRGYAGVAGYGTAFPGDGAGMLQGGPWWARIAWAV 120
QY 915 SOSSEEEQBEARAEASQSEEQBARAEPLQVSAAPVPEVGRAPTRSPPTPMDICQV 974
Db 121 SOSSEEEQBEARAEASQSEEQBARAEPLQVSAAPVPEVGRAPTRSPPTPMDICQV 180
QY 975 SLVQIRDLSDGDAADTSLSDISEVDPAVLNLSLDYDIKYLFFEFMI FRKVPKSAQPEPP 1034
Db 181 SLVQIRDLSDGDAADTSLSDISEVDPAVLNLSLDYDIKYLFFEFMI FRKVPKSAQPEPP 240
QY 1035 SPMAEELAEFPPEPTWMPGELGPHAGLEITESESDVDALLAEAAVGRKRWKSSPSRLF 1094
Db 241 SPMAEELAEFPPEPTWMPGELGPHAGLEITESESDVDALLAEAAVGRKRWKSSPSRLF 300
QY 1095 HFGRLPLDEPAELGLRERVKASVEHISRLKGRPEGLEKSGPPKPKGLASFLSLGLK 1154
Db 301 HFGRLPLDEPAELGLRERVKASVEHISRLKGRPEGLEKSGPPKPKGLASFLSLGLK 360
QY 1155 SWDRAPTFRLBELSDETVLQGSVTLACQVSAQPAQAATWSKOGAPLESRRVLSATLKN 1214
Db 361 SWDRAPTFRLBELSDETVLQGSVTLACQVSAQPAQAATWSKOGAPLESRRVLSATLKN 420
QY 1215 FQLLTILVVAEDLGVYTCVSNALGTVTGVLKAEPPSSPCPDIGEVYADGVLLVM 1274
Db 421 FQLLTILVVAEDLGVYTCVSNALGTVTGVLKAEPPSSPCPDIGEVYADGVLLVM 480
QY 1275 KPVSYPVTYIIVQCSLGGSWTTLASDIFDCCYLTSKLGRGTYTFTACVSKAGMGPY 1334
Db 481 KPVSYPVTYIIVQCSLGGSWTTLASDIFDCCYLTSKLGRGTYTFTACVSKAGMGPY 540
QY 1335 SSPSEQVLLGGPSHLASEEESQGRSAQPLPSTKTFQIQRGRFVSVRQCEKASGRA 1394
Db 541 SSPSEQVLLGGPSHLASEEESQGRSAQPLPSTKTFQIQRGRFVSVRQCEKASGRA 600
QY 1395 LAAKIIPYHPKDKTAVLREYALKGRHPHLAQLHAAYLSRHLVLIILELCSGPELLPCL 1454
Db 601 LAAKIIPYHPKDKTAVLREYALKGRHPHLAQLHAAYLSRHLVLIILELCSGPELLPCL 660
QY 1455 AERASYSSEVKDYLMQMLSATQYLHNOHILHDLRSENMIITEYNLLKVVLDLGNASLS 1514
Db 661 AERASYSSEVKDYLMQMLSATQYLHNOHILHDLRSENMIITEYNLLKVVLDLGNASLS 720
QY 1515 QEKVLPDSKDKYLETMAPELLEGQAVPQTDIWAIGVTAFLMIAEYFVSSEKARDLQR 1574
Db 721 QEKVLPDSKDKYLETMAPELLEGQAVPQTDIWAIGVTAFLMIAEYFVSSEKARDLQR 780
QY 1575 GLRKGVLRLSRVYAGLSGGAFLRSTLCAQPWGRPCASSCLOCPWLTTEGPACSRPAPV 1634
Db 781 GLRKGVLRLSRVYAGLSGGAFLRSTLCAQPWGRPCASSCLOCPWLTTEGPACSRPAPV 840
QY 1635 TPTARLVRVFNREKRALLYKRNHLAQR 1665
Db 841 TPTARLVRVFNREKRALLYKRNHLAQR 871

RESULT 12

US-10-307-019-8
; Sequence 8, Application US/10307019
; Publication No. US20030108533A1
; GENERAL INFORMATION:
; APPLICANT: Zeng, Wenlin
; APPLICANT: Stanton, Lawrence
; APPLICANT: SCIOS, INC.
; TITLE OF INVENTION: MEDIATORS OF SIGNAL TRANSDUCTION
; FILE REFERENCE: SCIOS.021DV1
; CURRENT APPLICATION NUMBER: US/10/307,019
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/548,473
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/129,552
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0

US-10-307-019-9

Query Match 27.6%; Score 2415.5; DB 14; Length 548;
Best Local Similarity 85.7%; Pred. No. 8.2e-113;
Matches 468; Conservative 25; Mismatches 52; Indels 1; Gaps 1;

1121 HISRIKGRPEGEKEGPPRRKPGIASFRLSGKSWDRAPTFRLSDETVVLGQSVTLA 1180
3 HISRIKGRPEGEKEGPPRRKPGIASFRLSGKSWDRAPTFRLSDETVVLGQSVTLA 62

1181 COVSAQPAQAATWSKDGAPLESSRVLSATLKNFOLLTILVVVAEDLGVVTCVSNALG 1240
63 COVLAQPAQAATWSKDGAPLESSRVLSATLKNFOLLTILVVVAEDLGVVTCVSNALG 122

1241 TVTTTGVLRKAERPPSSPCPDIGEVYADGVLLVWKPVEGYPVTYIVQCSLEGGSWTLA 1300
123 TAVTTGVLRKAERPPSSPCPDIGEVYADGVLLVWKPVEGYPVTYIVQCSLEGGSWTLA 182

1301 SDIFDCCYLTSKLSRGTYTFTACVSKAGMGPYSSPSEQVLLGGPNSHLASEEB-SOGRS 1359
183 SDISDCCYLTSKLSRGTYTFTACVSKAGMGPYSSPSEQVLLGGPNSHLASEEB-SOGRP 242

1360 AQLPSTKTFAQTQIRGRFSVVRQCKEASGRALAAKIIPYHPKDKTAVLREYALKG 1419
243 AQLPSTKTFAQTQIRGRFSVVRQCKEASGRALAAKIIPYHPKDKTAVLREYALKR 302

1420 LRHPHQAQLHAAYLSPRHLVLELCSGPPELLPCLAEASYSSEVSKDYLMWMLSATQYL 1479
303 LRHPHQAQLHAAYLSPRHLVLELCSGPPELLPCLAEASYSSEVSKDYLMWMLSATQYL 362

1480 HNQHILHDLRSNMILTEYNLLKVDLGNAGSLSQSKVLPSPDKFYLETMAPELLEGG 1539
363 HAQHILHDLRSNMILTEYNLLKVDLGNAGSLSQSKVLPSPDKFYLETMAPELLEGG 422

1540 GAVPQTDIWAIGTAFIMLSAEYPSVSEGDRLQGLRGLVRLSRCVAGLSGGAVAPLR 1599
423 GAVPQTDIWAIGTAFIMLSAEYPSVSEGDRLQGLRGLVRLSRCVAGLSGGAVAPLR 482

1600 STLCAQPWGRPCASSCLOCPMTTEGPACSRPAPVTFPTARLRFVVRNKRALLYKXH 1659
483 STLCAQPWGRPCASSCLOCPMTTEGPACSRPAPVTFPTARLRFVVRNKRALLYKXH 542

1660 NLAQVR 1665
543 NLAQVR 548

RESULT 14

US-10-425-114-37530
; Sequence 37530, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-136-D3_FLI.pap
US-10-425-114-37530

US-10-307-019-8

Query Match 28.0%; Score 2447.5; DB 14; Length 548;
Best Local Similarity 86.4%; Pred. No. 2.1e-114;
Matches 472; Conservative 27; Mismatches 46; Indels 1; Gaps 1;

1121 HISRIKGRPEGEKEGPPRRKPGIASFRLSGKSWDRAPTFRLSDETVVLGQSVTLA 1180
3 HISRIKGRPEGEKEGPPRRKPGIASFRLSGKSWDRAPTFRLSDETVVLGQSVTLA 62

1181 COVSAQPAQAATWSKDGAPLESSRVLSATLKNFOLLTILVVVAEDLGVVTCVSNALG 1240
63 COVLAQPAQAATWSKDGAPLESSRVLSATLKNFOLLTILVVVAEDLGVVTCVSNALG 122

1241 TVTTTGVLRKAERPPSSPCPDIGEVYADGVLLVWKPVEGYPVTYIVQCSLEGGSWTLA 1300
123 TAVTTGVLRKAERPPSSPCPDIGEVYADGVLLVWKPVEGYPVTYIVQCSLEGGSWTLA 182

1301 SDIFDCCYLTSKLSRGTYTFTACVSKAGMGPYSSPSEQVLLGGPNSHLASEEB-SOGRS 1359
183 SDISDCCYLTSKLSRGTYTFTACVSKAGMGPYSSPSEQVLLGGPNSHLASEEB-SOGRP 242

1360 AQLPSTKTFAQTQIRGRFSVVRQCKEASGRALAAKIIPYHPKDKTAVLREYALKG 1419
243 AQLPSTKTFAQTQIRGRFSVVRQCKEASGRALAAKIIPYHPKDKTAVLREYALKR 302

1420 LRHPHQAQLHAAYLSPRHLVLELCSGPPELLPCLAEASYSSEVSKDYLMWMLSATQYL 1479
303 LRHPHQAQLHAAYLSPRHLVLELCSGPPELLPCLAEASYSSEVSKDYLMWMLSATQYL 362

1480 HNQHILHDLRSNMILTEYNLLKVDLGNAGSLSQSKVLPSPDKFYLETMAPELLEGG 1539
363 HAQHILHDLRSNMILTEYNLLKVDLGNAGSLSQSKVLPSPDKFYLETMAPELLEGG 422

1540 GAVPQTDIWAIGTAFIMLSAEYPSVSEGDRLQGLRGLVRLSRCVAGLSGGAVAPLR 1599
423 GAVPQTDIWAIGTAFIMLSAEYPSVSEGDRLQGLRGLVRLSRCVAGLSGGAVAPLR 482

1600 STLCAQPWGRPCASSCLOCPMTTEGPACSRPAPVTFPTARLRFVVRNKRALLYKXH 1659
483 STLCAQPWGRPCASSCLOCPMTTEGPACSRPAPVTFPTARLRFVVRNKRALLYKXH 542

1660 NLAQVR 1665
543 NLAQVR 548

RESULT 13

US-10-307-019-9

; Sequence 9, Application US/10307019
; Publication No. US20030108533A1
; GENERAL INFORMATION:
; APPLICANT: Zeng, Wenlin
; APPLICANT: Stanton, Lawrence
; APPLICANT: SCIOS, INC.
; TITLE OF INVENTION: MEDIATORS OF SIGNAL TRANSDUCTION
; FILE REFERENCE: SCIOS.021DV1
; CURRENT FILING DATE: 2002-11-26
; PRIOR FILING DATE: 2000-04-13
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Mus musculus

Best Local Similarity 100.0%; Pred. No. 1.8e-94;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1276 PVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGTTTFTACVSKAGMGPYS 1335
Db 1 PVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRGTTTFTACVSKAGMGPYS 60
QY 1336 SPSEVILGGPSHLASEESQGRSAQPLPSTKTFATQIQGRFVSVRQWEKASGRAL 1395
Db 61 SPSEVILGGPSHLASEESQGRSAQPLPSTKTFATQIQGRFVSVRQWEKASGRAL 120
QY 1396 AAKIIPYHPKDTAVLREYALKGRHPHLAQLHAAYLSPHLVLIILELCSGPPELLPCLA 1455
Db 121 AAKIIPYHPKDTAVLREYALKGRHPHLAQLHAAYLSPHLVLIILELCSGPPELLPCLA 180
QY 1456 ERASVSESEVKQMLSATQYLNQHLHLDRSENMIITEYNLLKVDLGNASLSQ 1515
Db 181 ERASVSESEVKQMLSATQYLNQHLHLDRSENMIITEYNLLKVDLGNASLSQ 240
QY 1516 EKVLPDSDFKDYLETWAPELLEGGQAVPQTDIWAIGTAFIMLSAEPVSSEGGARDLQRG 1575
Db 241 EKVLPDSDFKDYLETWAPELLEGGQAVPQTDIWAIGTAFIMLSAEPVSSEGGARDLQRG 300
QY 1576 LRKGLVRLSRVAGLSGGAVAFRLSTLCAQPMGRPCASSCLQCPWLTEEGPACSRPAPT 1635
Db 301 LRKGLVRLSRVAGLSGGAVAFRLSTLCAQPMGRPCASSCLQCPWLTEEGPACSRPAPT 360
QY 1636 FTALRIVFVRNREKRALLYKRNLAQVR 1665
Db 361 FTALRIVFVRNREKRALLYKRNLAQVR 390

RESULT 15

US-10-333-314-18
; Sequence 18, Application US/10333314
; Publication No. US20030211093A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: KHAN, Farrah A.; GURURAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.; CHAWLA, Narinder K.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: GANDHI, Ameena R.; POLICKY, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.
; APPLICANT: THORNTON, Michael B.; LIU, Yan
; APPLICANT: NGUYEN, Daniel B.; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; LAL, Preeti G.
; APPLICANT: DING, Li; YAO, Monique G.
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.
; APPLICANT: KEARNEY, Liam; LU, Dyung Aina M.
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom
; APPLICANT: XU, Yuming; WALSH, Roderick T.
; APPLICANT: GIETZEN, Kimberly J.; YANG, Junming
; APPLICANT: JACKSON, Jennifer L.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0162 USN
; CURRENT APPLICATION NUMBER: US/10/333,314
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/23092
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/220,038
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/222,112
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,831
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,729
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 2380
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7477141CD1
US-10-333-314-18

Query Match 22.6%; Score 1973; DB 12; Length 2380;
Best Local Similarity 29.3%; Pred. No. 6.4e-90;
Matches 568; Conservative 236; Mismatches 638; Indels 498; Gaps 50;

QY 54 PSMQVITDVQAGTGAFAIIEGDPQPSVTWYKDSVOLVDSTLSQQQGTTSVLV 113
Db 598 PPSIMEDVEVGAGTARFAVVEGKPLFDIMWYKDEVLLTSSHSVSVYEENECSLVV 657
QY 114 RHVASDAGVYTCIAQNTGGVLCFAELVLGSDN---EPDSEKQSHR-RKLHSEYEVK 168
Db 658 LSTGADGGVYTTCTAQLAGEVSCAEALAVHSAQTAHMEVGEVGEDDHGRRLSDYDTH 717
QY 169 EETGRGVFGVFKVQHKMKILCAAKFIFLRGRTRAQAYRERDILAAALSHPLVTGLDQF 228
Db 718 QEIGRGAFAVYLRIVERSGLEFAAKFIPSAQPKASARREARLLARLOCHDCVLYPHEAF 777
QY 229 ETRKTLILILELCSSEHLLDRLYRKGVVTEAEVKVYIQQVLEGLHSHGVHLHDKPS 288
Db 778 ERRRGVIVITELC-TEELLERJARKPTVCESEIRAYMRQVLEGIHYLHSHVHLHDKPE 836
QY 289 NILMVHPA--REDIKICDFGFAQNIITPAELQFSQSGSPFVSPETIQONPVSEASDIWAM 346
Db 837 NLLVMDGAAGEQQVRCIDFGNAQELTGPBPQYCOYGTPEFVAPEIVNQSPVSGVTDIWPV 896
QY 347 GVISYLSLTCSSPAGESDRATILNVLEGRVSNSSPMAHLSSEDAKDF-IKATLQAPQA 405
Db 897 GVVAFLCTGISPFVGENDRITLLMNRNVNVAPEETFTLSLREARGFLIKVLVQ--DRL 954
QY 406 RPSAAOCLSHPWFLKSMPEAEAHFINTKQLFLLARSRWQSRMSYKSLVWMSIPELLR 465
Db 955 RPTAETLEHPFKTOAKGAB---VSTDHLKFLSRRWQSRQISYKCHLVLPPELLR 1011
QY 466 GPPDSPSLGVARHLCRDTGGSSSSSSSSONEL-----APFARAK-SLPPSPVTH 513
Db 1012 APPERVVMTMPRR-PPPSGGLSSSDSEEELEELPSVPRPQLOPFGSGRVSITDPTED 1070
QY 514 SPILHPRGFLRPSASLPEEAESERSTEAPAPASPEGAGPAAQGCVPFRHSVIR----- 568
Db 1071 EALGTETGATPMDWQEQGRAPSOQEAAPSFPALSPQGEPA--GASPRGELRGSSA 1129
QY 569 -----SLFVHQAG--ESPHEGALAPG-----SRHPARRHLLK 600
Db 1130 ESALPRAGPRELGRGLHKAASVELPQRRSPGPGATRLARGGLGEGBEYAOQLQRLRLR 1189
QY 601 GGYIAGALPGLREPLMEH-----RVLEEEAAREEQATL---LAKAPSFETALR 645
Db 1190 GGPEDGVSLRGLRFLLESIGGRARDPRMAAASSEAPHPQPPLENRGLQKSSFSQGEA 1249
QY 646 LPASGTHLAFGHSHSLEHDSPTPR-----PSSACGEAQRLPSAGGAPIRDMGHPOQS 701
Db 1250 EP-RGRHRRAGAPLETPVARLGARLLQESPSLSALSEAQ--PSSPA-----RPSAP 1297
QY 702 KQLPSTGGHPTAQPERPSPDSEWGPAPFCHPKQGSAPQEGCSPHPAVAPCPGGFPFG 761
Db 1298 K--PST---PKSAEPSSATPSDAPQPPAP--QAQKAPPEPRPEPVRAKPEAP---PPQ 1346
QY 762 SCK--EAPLVPSPPFLGQ-----POAPPAPAK-----ASPPLDSKMG 797
Db 1347 ALQTLALPLTFYAIQSLQSLGHAQPSQGPAAPEPKPHAAVFAVVASPP----- 1399
QY 798 GDLSLPGRPKPGPCSPGASQSSQVSLRVGSSQVGTETPGPSLDAEGWTOEAEPLSD 857
Db 1400 -----PGAEKRVPSAGGPPVLAEKARVPT-----VPPRPGSSLSSTSENLESEAVFE 1447
QY 858 STPTLQSPQEVMTWRFSLGGRGGYAGVGTGFAFGDAGGMLGQGPMMWARIWAVSQS 917
Db 1448 A--KFKASRES-----PLSLGLR-----LLSRS 1468

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2004, 15:05:32 ; Search time 28 Seconds
(without alignments)
5719.960 Million cell updates/sec

Title: US-10-697-263-2
Perfect score: 8740
Sequence: 1 MGCCRLGCGCSVAHSVSG.....RNREKRALLYKRNHQAQRV 1665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Piri:*
2: Piri2:*
3: Piri3:*
4: Piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	702.5	8.0	1398	2 T25568	hypothetical prote
2	550.5	6.3	1176	2 JN0583	myosin-light-chain
3	544.5	6.2	1147	2 A59307	myosin-light-chain
4	532	6.1	1906	1 S68235	myosin-light-chain
5	531	6.1	6658	2 T13931	projectin fruit
6	523.5	6.0	2783	2 T34416	hypothetical prote
7	498	5.7	6831	2 A88852	protein unc-22 (im
8	498	5.7	6839	2 S57242	twitchin (similar
9	498	5.7	7160	2 T27935	hypothetical prote
10	465.5	5.3	608	2 A35021	myosin-light-chain
11	463.5	5.3	451	2 S49128	twitchin-like prot
12	456	5.2	1423	1 I37275	death-associated p
13	454.5	5.2	1398	2 T13741	hypothetical prote
14	451	5.2	371	2 JC7733	death associated p
15	450	5.1	610	2 A28798	myosin-light-chain
16	449	5.1	1051	1 JW0051	serine/threonine-s
17	442.5	5.1	26926	1 I38344	titin, cardiac mus
18	441.5	5.1	374	1 S50193	Ca2+/calmodulin-de
19	438	5.0	1734	2 A54502	microtubule-associ
20	427.5	4.9	1435	2 T32930	hypothetical prote
21	427	4.9	370	1 S57347	Ca2+/calmodulin-de
22	423.5	4.8	1211	2 T27522	hypothetical prote
23	418	4.8	414	2 JN0323	Ca2+/calmodulin-de
24	407.5	4.7	606	2 D96532	probable CDPK-rela
25	406	4.6	602	2 S60052	calcium-dependent
26	404.5	4.6	469	1 S17656	Ca2+/calmodulin-de
27	401.5	4.6	601	2 T46084	CDPK-related prote
28	401	4.6	529	1 S71774	calcium-dependent
29	399.5	4.6	473	1 A53036	Ca2+/calmodulin-de

30 398.5 4.6 708 2 T23616 hypothetical prote
31 398.5 4.6 1237 2 T45070 protein kinase hom
32 398 4.6 1192 2 T18611 probable serine/th
33 394.5 4.5 348 2 T37321 Ca2+/calmodulin-de
34 391.5 4.5 608 2 T18445 hypothetical prote
35 389.5 4.5 474 1 TVRTCA Ca2+/calmodulin-de
36 389.5 4.5 502 2 I52637 Ca2+/calmodulin-de
37 387 4.4 533 2 C32571 ribosomal protein
38 387 4.4 740 2 I38556 ribosomal protein
39 386.5 4.4 560 2 S51600 phosphorylase kina
40 385 4.4 752 1 A32571 ribosomal protein
41 384.5 4.4 774 2 I48609 probable serine/th
42 384 4.4 735 2 A53300 ribosomal protein
43 383 4.4 481 2 JEO377 p70 S6 kinase (EC
44 381.5 4.4 888 2 A55318 serine/threonine p
45 381.5 4.4 1246 2 G89287 protein H39223.1 [

ALIGNMENTS

RESULT 1

T25568
hypothetical protein C24G7.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25568
R;Goela, D.; Wilson, R.

A;Description: The sequence of C. elegans cosmid C24G7.
submitted to the EMBL Data Library, February 1997

A;Reference number: Z20052

A;Accession: T25568

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1398 <GOS>

A;Cross-references: EMBL:U88310; PIDN:AAB42336.1; GSPDB:GN00019; CESP:C24G7.5
A;Experimental source: strain Bristol N2; clone C24G7

C;Genetics:

A;Gene: CESP:C24G7.5

A;Map position: 1

A;Introns: 13/2; 132/2; 220/3; 316/2; 357/1; 491/3; 558/3; 734/3; 842/1; 1103/3; 1213/2;

Query Match 8.0%; Score 702.5; DB 2; Length 1398;
Best Local Similarity 22.2%; Pred. No. 1.4e-15;
Matches 366; Conservative 215; Mismatches 582; Indels 485; Gaps 70;

QY	204	AQAYRR	-----DILAALSHPLVTGLDQPETRKTLILILELCSBELLDRLYRK	253
DB	2	AEQYEQKQRTLYANSTIDGLSLAHPGV	-----EIAE---PK	35
QY	254	GVVTEAEVKYIQQLVEGLHYLHSHGVHLDIKPSNIMLVHPAREDIKICFGPAQNI	TP	313
DB	36	GVNRETQVRFVQRLALAKXMDLRIAHLDLRPETILL	---QDDKLADFGQARRLLR	92
QY	314	AELOFSQSGSPFVSPFETIQQNVSEASDIWAMGVISYLSLTCSPPAGESDRLTNVL		373
DB	93	GLITGEIKSGPFVSPFVSVPLATDMSTGVLTVTLLTGLSPHGDNDNETLANVD		152
QY	374	EGRVNSSMAHLSEDAKDFIKATLQAPQAPSAOCLSHPW	-----LKSMAEAEHF	429
DB	153	SCQFD--SSEL--GNFSYDAGDFVKLLTIPVSRULTVDALDHPWINDKLTETP	-----	204
QY	430	INTKQLKFLARSQRQSLMSYKSLVMRS	-----IPELLRGP	479
DB	205	LSADTLREFKYCHKLE	-----RRVFVQQTPEQILEAILGPATAQAQNAFVAPEGRRP	259
QY	480	CR	-----DTGSSSSSSSSDNE	504
DB	260	AIYDYLRIQPKXPPTVEYVQPKRKHPPIDBFGLQIDGADFREGTGFEGPHQPP		319
QY	505	SLPPSP	-----VTHS	535
DB	320	QIPPPQRENQAADSRRAHQEQPQHQGPQIPVDQYGRPLVDPR	-----YLNDFSHRPSLDD	378

QY 536 SERSTEAPAPSPAGGAPAAQCVPRHVSIVISLFVHOAGBSPHGHALAPGSRBHPARR 595
DB 379 AFFVYKXGNVPHFDKXGRMA-----PQMLEKKLIPQDKGTSPHS--KKEKTQHPVAT 432
QY 596 RHLKGGYIAGALPG-----LREPLMEHVLBEBAAR-----BEQATLLAKAPSP 640
DB 433 PIL-----ASFGSDQOQKIPMRMIRGERIEEIEIANRILSDISEGSIAGSLASL 484
QY 641 ETALRLPASGTHLAPGHSHLEHDSPTTPRPSSEACGEAQLRPSAPSGGAPIRDNGHPQG 700
DB 485 ED-PEIP-----KDFQVEASEPSTPLTPE-----VIRE-----513
QY 701 SKQLPSTGGHPTAGAPERPSPDPSWGPQAPFCHFKGSGAPQEGCSPHPAVAPCPGSPFP 760
DB 514 -----TIKPTPTSTPTSPKSPVP-----531
QY 761 GSCKEAPLVPSPPLGQOAPAPAKAS-----PPLDSKM-----GPGDISLP-GRP- 806
DB 532 -----QPQGLLPVAKVTVSDSILAGLPAADKKVLEDAENDPSIFVGAPL 575
QY 807 -KPGPCSPGSAQASQSSQVSSLRVGSQVGTGPGSLDAEGWTEAEDLSDSTTLQRP 865
DB 576 FLEGLHSGDITDITTSAGSL--IKVTSAINLSPNPK-----SPRSTPGTKSPVWLSP 627
QY 866 QEQVTKRKFSLGGRGGVAGVAGYGTAFAGDAGMGLQGPMWARIANAVSQSEEEEOEEA 925
DB 628 RQEHSMELVATKRG-----XEGFLPPGEL-----AEDIDDEDA 661
QY 926 RAESQSEGEQAEASPLPOVSARPVDEVRAPTRSSPEPTPWEIDIGQVSLVQIDRLSGD 985
DB 662 FMDRKKQ-----VKPDHDE 678
QY 986 ABAAD-----TISL-DISEVDA--YLNLSL-----YDKVLPPE--PMI--- 1021
DB 679 NDFKEXERLEKOKNRVTNLDLKVPSAFYKQDSDFGHGYDIDATPNDSHVQIGPD 738
QY 1022 -FRKVPKSAQPEPPSMAEBELAEFPEPTWPGELG-PHAGLETESESDVALLAA 1079
DB 739 TYLMAAGAAFNKRVNREBELFGMGAPVK-QGFLGVNRRDITVRERRYTD-ILRETT 796
QY 1080 VGRKRKWSFSLFHPGRHLPLDEPAELGLRERVKASVEHISILKGRPEGLEKEGPP 1139
DB 797 QGLEPKSHEQSTAL-----LOKAPSATALEIRIKADIEKVTCP-----A 834
QY 1140 RKXGLASFRSLGLKSHDRAPTFLRELSDETIVLQGSVTLAQVQAQAAQATWSKQAP 1199
DB 835 TKKNDGTF-----APIFTARLDVYLRKNQPAIFCAVASAPAPKVTFDFQKI 884
QY 1200 LESSRYLISATLKNFOLLIVVAEDLGVYTCVSNALGTVTITGVLRKAERPSSSPC 1259
DB 885 LESNDRTVEQD--NNVARLILNHAAPYDLGEYVCTAINEYGTDXKSCRLISGETSPRGR 943
QY 1260 PDICEVYAD-GVLLWKPVSYP-----VTIIVQCSL-----EGSWTTLASDIFDCY 1308
DB 944 PE-AELSSDTEIFQWEAPB--GFTYLEGITYLEVRVAGPNHDGDPMTVSEKIDDES 1000
QY 1309 LTSKLSRGGYTTRTACVSKAGMPYS-----SPSEQV-----1341
DB 1001 IVKHLSPGIIYQFRVTAQNGFGLGLPSLSRIIVQTHGKGAAPKQIDVLKSEIRLNVSM 1060
QY 1342 -----ILGGSHLASEESQRA-----OPLPSTKTPAQTOIQGRPSVWQC 1386
DB 1061 QKSTNQLGGSSE-ESEEDSEARTANEMKSNLOQDPTGTFQGLGFKGRFVSIRDA 1119
QY 1387 WEKAS-GRALAAKIIPYHKDITAVLREYALKGLRPHPLAQLHAAYLSPRHLVILEL 1444
DB 1120 VDS*TEGHACHAVKI--RHPSSB--AISEYSLRDQGHENVQRLIAAFNNSFLYLSER 1175
QY 1445 CSGPELLAPCLARASISESVKYLWQMLSATQYLNQHLHLDLRSENMIITEYN--LL 1502
DB 1176 LY-EDVFSRFRVENDYITEEQVALTMQVTSALHFLHFKGLAHLVDNPHNMFQSKRSWV 1234

QY 1503 KVVLDLNAQSLQSKVLPSPDKFDYLET--MAPEL-LEGQGAVPOTDIWAIGVTAIFMLS 1559
DB 1235 KLVDFOGAQKVS-GAVKPD-----PDKWASPEPHETPVTQSDMWGKGVVTFCLLA 1288
QY 1560 AEPVSSSEGARDLQRLKGLVRLSRCYAGL-----SGGAVAFILRSTICAPGWRPCASS 1614
DB 1289 GFHPPTSE--YDRBEIEKENVIN-KCDPNLIPVNASQECLSFATWALKKSPVRMRTRDE 1345
QY 1615 CLQCPWLTEEGPACGRPAFVTPPTARLR 1642
DB 1346 ALSHKFLSSDPSMVRRESIKYSASELR 1373
RESULT 2
JN0583
myosin-light-chain kinase (EC 2.7.1.117) 155K protein - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Apr-2000
C:Accession: JN0583; P0490
R:Kobayashi, H.; Inoue, A.; Mikawa, T.; Kuwayama, H.; Hotta, Y.; Maeaki, T.; Ebashi, S.
J. Biochem. 112, 786-791, 1992
A:Title: Isolation of cDNA for bovine stomach 155kDa protein exhibiting myosin light ch
A:Reference number: JN0583; MUID:93203148; PMID:1284247
A:Accession: JN0583
A:Molecule type: mRNA
A:Residues: 1-1176 <KOB>
A:Cross-references: GB:S57131; NID:g298638; PIDN:AB25794.1; PID:g298639
A:Experimental source: stomach
A:Accession: P0490
A:Molecule type: protein
A:Residues: 44-35;721-728;828-851;1002-1019 <K02>
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; fibronectin type III
C:Keywords: ATP; calmodulin binding; phosphotransferase
F:370-430/Domain: immunoglobulin homology <IMM1>
F:723-980/Domain: protein kinase homology <KIN>
F:731-739/Region: protein kinase ATP-binding motif
F:1083-1144/Domain: immunoglobulin homology <IMM2>
Query Match 6.3%; Score 550.5; DB 2; Length 1176;
Best Local Similarity 25.3%; Pred. No. 7.9e-11;
Matches 145; Conservative 91; Mismatches 191; Indels 147; Gaps 10;
QY 45 SSLPALPG-----PPSMQVITIEDVQAQTGGTAQFEAIIEGDPQP 83
DB 468 SSIPVLTGSDATVKKPAPKTPPKAAMPQIIQPEDQKVRAGESVLEFGKVGATQPI 527
QY 84 SVTWKDSVQLVDSTRLSQQGGTTSYLVLRHVASKDAGVITCLAQN-----130
DB 528 TCTWMKFRQIQDSBHIKVENSEQSGSKLTIRAAQEHGCGCYLLVENKLGSRQAVNLTV 587
QY 131 -----TGGQVL-----CK-- 138
DB 588 VDKPDPACTPCASDIRSSSLTSLVSGSYDGGSAVQSVSVIWDSDVKTWKLATCRST 647
QY 139 -----AELLVLG--GDNEP-----DSEKQ-- 155
DB 648 SFNVQDLLPDREYKFRVRAINVYGTSPSQSESLTALGEKPEEPFKDVEVSDDEKEPE 707
QY 156 -----SHRKLHSFYVEKEEIGRGVFGFVKRVQHKGNKILCAAKFIPLRSTRQAQY 207
DB 708 VDYRTVTVNTEQKVSDFYDIEERLGGKFGQVFLVRLEKTKGKIWAGKFFKAYSAREN 767
QY 208 R-ERDILALSPVLTGLLDQETKTLILILELCSSELDRLYRKGV-VTEAEVKYI 265
DB 768 ROEISIMNCHHPKLVQCVDPAFEKANIMVLEIVSGGELFERIDEDFELTERECIKYM 827
QY 266 QQLVEGLHYLHSHGVHLDIRKPSNIMVHPAREDIKICDFGAQNTTFAELQFSQYSPE 325
DB 828 KQISEGVEYIHKQGIIVHLDLKPNIMCVNKTGTRIKLIDFGLARRLENAGSLKVLFGTPE 887
QY 326 FVSPETIQNPVSEASDIWAKGVISYLSLTSSPAGESDRATLNLVLEGRVSSPMAA 385
DB 888 FVAPEVINTPEIGYATDMWSIGVICIILVGLSPFMGNDNETLANVTISATWDFDDEAFD 947

QY 386 HLESDAKRTKATLQAPAPSAOCLSHPNFLKSPAEBAHFINTYKQLKELLARSRWQ 445
Db 948 EISDDAKDFISNLKDKMKNRLNCTQCLQHPWLMKDTQRMEXAKLSKDRMKYMARRWQ 1007
QY 446 RSMYSKSYSLVMSRPEL-----LRGPPDSP 471
Db 1008 KTGNAVRAIGRLSSAMISGLSRKSSSTGSP 1041
RESULT 3
A59307
myosin-light-chain kinase (EC 2.7.1.117), smooth muscle - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 20-Jun-2000 #sequence revision 20-Jun-2000 #text_change 20-Jun-2000
C/Accession: A59307; A41674; B41674; A40210
R:Gallagher, P.J.; Herring, B.P.; Griffin, S.A.; Stull, J.T.
J. Biol. Chem. 266, 23936-23944, 1991
A>Title: Molecular characterization of a mammalian smooth muscle myosin light chain kinase
A/Reference number: A41674; MUID:192084694; PMID:1748666
A/Accession: A59307
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-1147 <GAL>
A/Cross-references: GB:M76233; NID:g165703; PIDN:AAA73093.1; PID:g165704
A/Experimental source: dev stage Adult; sex female; tissue type smooth muscle
A/Note: this revision submission is not cited in GenBank entry RABSMMLCKR, release 115
A/Accession: A41674
A/Molecule type: DNA
A/Residues: 1-22, 'CTA', 26, 'SRS', <GA>
A/Cross-references: GB:M76369
A/Note: this translation is not annotated in GenBank entry RABSMMLCKD, release 115
A/Note: this sequence has been revised in reference A40210
R:Gallagher, P.J.; Herring, B.P.; Griffin, S.A.; Stull, J.T.
J. Biol. Chem. 267, 9450, 1992
A/Reference number: A40210; MUID:92250555; PMID:1577772
A/Contents: erratum
A/Accession: A40210
A/Molecule type: DNA
A/Residues: 23-30 <GA>
A/Note: this is a revision to the sequence B41674 from reference A41674
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; fibronectin type III
C/Keywords: ATP; phosphotransferase; smooth muscle
F:343-403/Domain: immunoglobulin homology <IMM>
F:694-951/Domain: protein kinase homology <KIN>
F:702-710/Region: protein kinase ATP-binding motif
F:1055-1116/Domain: immunoglobulin homology <IMM2>
Query Match 6.2%; Score 544.5; DB 2; Length 1147;
Best Local Similarity 23.7%; Pred. No. 1.2e-10;
Matches 149; Conservative 97; Mismatches 206; Indels 177; Gaps 10;
QY 48 PALGPPSPQVTEIDVQAGTGAQFEALISGDPPSVTWYKDSVOLVDSTRLSQQEGT 107
Db 463 PRAAPPQIPQPPQDQKVRAGVERELFGKVGATGPICTWKKFRKIQDSEHKVNSEA 522
QY 108 TVSLVLRHVASKAGVYTCIAQTGQVCLKAEILLVLG----- 145
Db 523 GSKLTLAAROEHCYGLLVENKLSRQAQVNLTVVDKPPDPAGTFCASDIRSSSLTSL 582
QY 146 -----G 146
Db 583 WYGSSVDGGSVQSVSVIWDSDVKMTLTCSTSNVLDLPPDREYKFRVRAINVG 642
QY 147 DNEP-----DSEKQ-----SHRKLHSFYVEKKEI 171
Db 643 TSEPSOESELTWVGKBPBPKEVEVSDDEKPEVDYRTVTVTNTEQKVSDFYDIEERL 702

QY 172 GRGVTFVYKVOHKGKILCAAKFIPLRSRTRQAQYRERD-----ILAALSHPLVTGL 224
Db 703 GSGKFGQVRLVEKKTGKIMAGKFF-----KAYSAREKENIPAEIGIMNCLHHPKLVOC 756
QY 225 LQOFTTRKTLILILBCSSBELDLRLVRKGV-VTEAEVKVYIQQLVEGLHYLHSGRVLHL 283
Db 757 VDAFEKANIYVWLEIVSGGELFERIIDDFELTERECIKYMRQISEGVEYHKGQIVHL 816
QY 284 DIKPSNIMLVHAPAREDIKICDGFPAQNTIPABLOFSQVSGSPFVSPILIQNPVSEASDI 343
Db 817 DLKPNIMCVNKTGTRIKIDIFGLARRLENAGSLKVLFGTPEFVAPEVINYEPIYSATDM 876
QY 344 WANGVISYLSLTSSPFFAGESDRATLNLVLEGRVSWSPMAAHLSEDAKDFIKATLQAP 403
Db 877 WSGIVICVILVSGLSPFFGNDNETLANVTSATWDFDEAFDEISDDAKDFISNLLKDM 936
QY 404 QAPSAACLSHPWFLKSPAEBAHFINTYKQLKELLASRQWSRSLMSYKSLVMSRPEL 463
Db 937 KRLDCTCLOLQHPWLMKDTKNWEAKKUSKDRMKYMARRWQKTGNVRAIGRUSSVAMI 996
QY 464 LRGPDPSPSLGVARHLCLRDTCGSSSSSSSSSDNELAPFARAKSLPPSPVTHSLHPRGFL 523
Db 997 -----SGLSRKSSSTGS-----PTSPLT----- 1014
QY 524 RPSASLPSEAESESTEAPA---PPASP 549
Db 1015 --AERLETEEDVSQAFLEAVAEKPHVKP 1041
RESULT 4
S68235
myosin-light-chain kinase (EC 2.7.1.117), 210K, nonmuscle - chicken
N:Contains: myosin-light-chain kinase, 108K, smooth muscle; telokin
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999
C/Accession: S68235; A43899; A44389; A44389; S28227; S78216; A35093; A25810; S11652
R:Watterson, D.M.; Collinge, M.; Lukas, T.J.; Van Eldik, L.J.; Birukov, K.G.; Stepanova, F.E.S. Lett. 373, 217-220, 1995
A/Title: Multiple gene products are produced from a novel protein kinase transcription re
A/Reference number: S68235; MUID:96033976; PMID:7589469
A/Accession: S68235
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1906 <WAT>
A/Cross-references: EMBL:X52876; NID:g992992; PIDN:CAA37056.1; PID:g992993
R:Shoemaker, M.O.; Lau, W.; Shattuck, R.L.; Kwiatkowski, A.P.; Matrisian, P.E.; Guerra-S
J. Cell Biol. 111, 1107-1125, 1990
A/Title: Use of DNA sequence and mutant analyses and antisense oligodeoxynucleotides to e
activity
A/Reference number: A37099; MUID:90361738; PMID:2202734
A/Accession: A37099
A/Molecule type: mRNA
A/Residues: 649-1906 <SHO>
A/Cross-references: EMBL:X52876
R:Collinge, M.; Matrisian, P.E.; Zimmer, W.E.; Shattuck, R.L.; Lukas, T.J.; Van Eldik, L.
Mol. Cell. Biol. 12, 2359-2371, 1992
A/Title: Structure and expression of a calcium-binding protein gene contained within a c
A/Reference number: A44389; MUID:92236611; PMID:1373815
A/Accession: A44389
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1750-1906 <CO2>
A/Cross-references: GB:M88283; NID:g211371; PIDN:AAA48647.1; PID:g211372
R:Yoshikai, S.; Ikebe, M.
Arch. Biochem. Biophys. 299, 242-247, 1992
A/Title: Molecular cloning of the chicken gizzard telokin gene and cDNA.
A/Reference number: S28227; MUID:93073972; PMID:1444462
A/Accession: S28227

QY 1331 MGYSSPSEQVLLG--GPHSLASBE---ESQGRSAQPLPSTKT-----FAFQTK 1374
Db 1399 ISEPSQSESVVKGKQBEELKEBEAELSDDEKETEYVYRTVTINTQKVSDVYNIER 1458
QY 1375 IQGRFSVROCKEAKSGRALAAKII--PYHPKOKTAVLREYALKGLRHPHQAOLHAAYL 1433
Db 1459 LGSKFGVGFVLEKTKVWAGFFXKAYSAREKENIRDEISINNCUHHPLVOCVDAFE 1518
QY 1434 SPRHLVILELCSGPELLP--CLABRASYSESEVDYLMQMLSATQYLNQHILHLDRSE 1492
Db 1519 EKANIMVLEWVSGELFERIIDFELTERECIKYMRQISEGVEYHKQGIHVHLDLKE 1578
QY 1493 NMITEY--NULLKVDVIGNAQLSQE---KVLPSDKFVDLETWAPELLESGQAVPOTDI 1547
Db 1579 NIMCVNKTGTSIKUIDFGLARLESAGSLKVLFGTP-----EFVAPEVINYEPIGYETDM 1633
QY 1548 WAIGVTAFIMLSABYPVSSEGCARDLQRLKGLVRL--SRCVAGLSGCAVAFRLSTLCAQP 1606
Db 1634 WSGIVCIVLVSGLSFPMGNDNETLANVTSATWDFDEAFDEISDDAKDFISNLLKKDM 1693
QY 1607 WGPCASSCLOCPWITEE 1624
Db 1694 KSLNCTOCLQHPWLQKD 1711
RESULT 5
Ti3931
projectin - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: Ti3931
R:Dailey J.; Southgate, R.; Ayne-Southgate, A.
J. Mol. Biol. 279, 201-210, 1998
A:Title: Structure of the Drosophila projectin protein: isoforms and implication for pro
A:Reference number: Z17815; MUID:98300339; PMID:9636710
A:Accession: Ti3931
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6658 <DAL>
A:Cross-references: EMBL:AF047475; NID:93337430; PID:93337431; PIDN:AAC27550.1
C:Genetics:
A:Gene: projectin
A:Cross-references: FlyBase:FBgn0005666
A:Map position: 4
A:Note: intron positions not resolved (incomplete sequence)
C:Keywords: muscle
Query Match 6.1%; Score 531; DB 2; Length 6658;
Best Local Similarity 20.7%; Pred. No. 1.6e-09;
Matches 277; Conservative 153; Mismatches 403; Indels 508; Gaps 46;
QY 158 REKLSFYEVEKEIGRGVGFVKRVQHGKNIKCAKFIPL-RGRTRAQAYRERDILAAL 216
Db 5707 QQSVYDRIDILEITGTGAFGVHRCRERSTGNIFAKFIPVSHSVEKDLIRREIDIMNQL 5766
QY 217 SHPLVTGLDQPETRKTILILELCSBELDLRYKRG-VVTEAFVKYIOQLVEGLHYL 275
Db 5767 HHOKLINLHDAFEDDDEMLILEFLSGGELPERITAEVGVVTEAEVINMRQICEGIRHM 5826
QY 276 HSHGVHLHDIKESNIMLVHPAREDKICDFQFAQNIITPAELQFSQYGSPEFVSPILQON 335
Db 5827 HEQNIHLHDIKESNIMLVHPAREDKICDFQFAQNIITPAELQFSQYGSPEFVSPILQON 335
QY 336 PYSEASDIWAMGVI SYLSITCSSPPAGESDRATLNLVLEGRVSWSPMAHLSEDAKDFI 395
Db 5887 PVGFYDMWATGVLSVYLLSGLSPFAGDNDQTLKNVACDWDFALESFKYISBEAKDFI 5946
QY 396 KATLQAPQARSACQLSHPHFLKMPABEAHFTINTQLKELLARSEWQSLSYKSIL 455
Db 5947 RXLLVRNEKRWTAHECLLHPWLTDGHSAMKQE--NRD--RYLAYREKLRRKYEDFERFL 6003
QY 456 V-----MRSIPELLRGPPDPSFGLVARHLCRDTGGSSSSSSSSSDNELAPFAKSLPP 508

Db 6004 LPIGRLSEYSSLRKLL-----MEKYKIH-----AVDRQAA-- 6036
QY 509 SPVTHSPLLHPGRFLRPGASLPPEEAASE--RSTEAPAPASPEGAGPPAAQGCV-- 561
Db 6037 -----PFVIRPSSQFCYEQSVKFCRCIAIATPTLTWNNHNIELRQSVKFMCR 6086
QY 562 -----PRHSVIRSLFTHQAGESPEHCALAPGSRHRPARRHLLKGGYIAGALPLREPLME 617
Db 6087 YVGDDYYFIINRVKPHDRGE-----YIIRA----- 6111
QY 618 HRVLEEAAREEQATLLAKAPSFETALRLPASGTHLAPGHSLSHLEHSPSTPRPSEACG 677
Db 6112 -----ENHYGSRREVFLNVQP-----LP-----KEQPRYRTES-- 6140
QY 678 EAQRLPSAPSGGAPTRDMGHPOGSKQLPSTGGHPGTAQPERSPSPGQAPAFCHPKQG 737
Db 6141 -----TPVR-----RREPLPYTFW----- 6154
QY 738 SAPOEGCSHPAVAPCCPGSPPGCKEAPLVSPSPFLGQOPAPAPAKASPLDSKMP 797
Db 6155 -----Q-----ESETAPSTFLLRPV-----MQA 6175
QY 798 GP-----ISLPGRPKPGPCSPGASQASQSSVSLRVSSQVGTPEPSPSLDABQWQBA 852
Db 6176 RDTCKLLCLSGKVP-----NVR-----WYKOG 6199
QY 853 EDLSDSPTPLQRPQVQVTKRPSLGRGGYAGVGTFAFGDAGGMLGQGPMMARIAW 912
Db 6200 RELSKYEYANTHSDGVVTFM----- 6219
QY 913 AVSSEEEQEEARAESEQEEQAEARABSPLPQVSGARPVEVGRAPTRSSPPTWEDIG 972
Db 6220 -----IIDCKP-----SDSG 6229
QY 973 QVSLVQIRLDGDAEADTISLDISEVDPAYNLSD--LY--DIKYLPEFWMIFRKVPKS 1028
Db 6230 NYS-CKATNCHGTDETCVVIIEGEMVTPVTEQQLAHNPLYSGDRKI-- 6275
QY 1029 AQPPEPSPMAEBELAEFFPTWMPGELGPHAGLITEESDVEDALLA---BAAVGRKK 1085
Db 6276 EQPIKAPLPVTSRQYTS-----SSVQNTSPQGDKNVNSNSGSGISNKKK 6323
QY 1086 WSPSPRSLSLPHFQGRHLPLDEPABLGLRERVKASVEHISRLKGRPEGLEKEGPPRKKPGL 1145
Db 6324 YASNS--LQAPG-----SFSRSRSAT 6342
QY 1146 ASPLSLGLKSWDRAPTFLRELSDETVVLGQSVTLACQVSAQAAQATWSKDGAPLESSSR 1205
Db 6343 KELLLPDDSLMCKPEFTKPLHDLIHDGEQILITCYVKGDPEPQISWSKNGKSLSSD- 6401
QY 1206 VLISATLKN-FQLLTILVVVADSLGVYTCVSNALGTVTTTGVL-----RKAER 1253
Db 6402 -ILDLYKNGIATLTINEVFPDEGEVITCTATNSVGAETKCKLTIQPLDKNINKKVA 6460
QY 1254 PSSSP-----CPDIGEVVADGVLVW-----KPVESYGPVTYI 1286
Db 6461 GDNAPKIVSHLSRFRVDRGDVNLACRIIQAQHF-----VWLHNNKEIKPSKDF----- 6511
QY 1287 VQSLGSGSWTTLASDIPCCVLTSLKSRGYT-----PFTACVSKAGMGPYS 1336
Db 6512 -QYTNANIVRLQIASIFP-----EDGGITVCEAFNDIGESFSTCTINVTVPDETK 6562
QY 1337 -----PSEQVLLGGPSHLASEESO-----GRSAQPLST-KTFAFQTIQGRFS 1381
Db 6563 QPSFVKPPTSVSVLEGEFTFCEIDISELLNLVWLKDGKPIDETLPRYFTKDGHRYSFA 6622
QY 1382 V-----VRQWEKA-SGRA 1394
Db 6623 VAKCNMDDVQYQAKAVSGKA 6643

734416
hypothetical protein F12F3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34416
R:Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A:Description: The sequence of C. elegans cosmid F12F3.
A:Reference number: Z21521
A:Accession: T34416
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2783 <FUL>
A:Cross-references: EMBL:U80022; PIDN:AAC25886.1; GSPDB:GN00023; CESP:F12F3.2
A:Experimental source: strain Bristol N2; clone F12F3
C:Genetics:
A:Gene: CESP:F12F3.2
A:Map position: 5
A:Introns: 45/3; 90/3; 451/3; 509/1; 2313/3; 2341/3; 2378/3; 2414/2; 2453/3; 2474/2; 252

Query Match 6.0%; Score 523.5; DB 2; Length 2783;
Best Local Similarity 20.6%; Pred. No. 1.2e-09;
Matches 364; Conservative 223; Mismatches 643; Indels 539; Gaps 73;

QY 48 PALGPPSMQ-VTIEDVQ-----AQTGTAQFAEIIIEGDPQPSVTWYKDSVQLVDSTRL 100
DB 50 PAAPGPAVEDQNVDSVRLWAAPTNDGGSPVRNVTVMCTEKGKWTAKAEVTKQAFITL 109
QY 101 SQOQEGTYSVLRHVASKDAGVYTCLAQNTGGQ-VLCKAELLY----- 143
DB 110 FNLVPGESYFRVR-----ADNTFGQSEPSDESELVVVKNVSVVVEPKKE 156
QY 144 -----LGCNPEPDEK--QSHR--RKLHSYEVKEE:GRGVFGFVRVQHK 186
DB 157 VKYKEQESVDYERVAKDSEPSYKTIIDHRLPND-QAKYIIEELGKAGYGTVYRATKA 216
QY 187 NKILCAAKFIPLRSRTAQ-AVRERDILAAISHPLVTGILDOFERKTLIILIELCSSEE 245
DB 217 TGTWAAKQVQVPGVKENVHEISMMNQLHHEKLLNHLAEAFDMGNEMWLLIEFVSGGE 276
QY 246 LLRLRYR-KGVTEAEVKYIQOLVEGLYHLSHGVLHLDIKPSNLMVHPAREDIKID 304
DB 277 LFKKILDDSLMBEEVRDYMHOILDGVSHMKHKNQIVHLDLKPENILLKAKNSNELKID 336
QY 305 FGPAQNTITPAELOFQSOYSPFVSPRIIOQNVPSEASDIWAMGVISYLSLTCSSPPFAGES 364
DB 337 FGLARKLDPKSVKLLFGT:PERCAPEVNVYQFVGLSTDWTVGVISYVLLSGLSPFLGDS 396
QY 365 DRATLNVLEGRVSNWSPMAHLSEDAKDFIKATLQAPQAPPSAAQCLSHPWFLKMPA 424
DB 397 DEDTLANVSASDWDFODPDSDVDSDLAQDFICRLMIKDKRMSVQDRLRHPWITKMQPK 456
QY 425 EEAHFTNKQLKPLLARSRWQSLMSYKSIIVMRGIPPELLRPPDPSLGVARHLCDTG 484
DB 457 LDKSGVPARQKRNFLSLKRWSDLLFGL-----AKRG-----AIFRRLTWD-- 499
QY 485 GSSSSSSSSDNELAPARAKSLP-----PSPVTH-----SPLLHP 519
DB 500 GVFERIAFTDAAPSVK-KQLEDIVANVGDIATLSCDVGDPSPKQVQYKDKELTVP 558
QY 520 RGLRPSASLPEAEAEASERSTAPAPSPAGGPPAAQ-----GCVPRHGVIRSLFVHQ 574
DB 559 -----SMKYDSFVNEGLAELTVKNIVESDAGYVTCRAINDLGSIMTEAKLSV---X 606
QY 575 AGSPHEGALAPQSRHPA--RRRLHLLKGYTAGALPGLRPLMEHVRVLEAEAREEQAT 632
DB 607 ADSKKK-----KSKSTSPAVIEKKDKRKTQKVVVTEIMDPPNFHLLQDDEAKIGEPK 661
QY 633 LLAKAPSFETALRLPA-----SGTHLAPGHSHL-EHDS-----PSTPRPSSEAC 676
DB 662 ILVVT---NTTLPPTVDVWHNGEHSINDSNYLAKHDKGRYELHILSVDSITDGKWKAV 718
QY 677 G-----EAQRLPSASGGAPADMDMGHPOGSKQLPSTGCHPQTAA----- 715

DB 719 GRNAFCECESEAKLTWVIPDG-----QYAPFGKQLSDVKCSSEDLILKEVN 765
QY 716 -PERSPSPWQAPPAFFCHPKQ-----GS-----APOGCGSPHAPVAPCPGGS 757
DB 765 IQANPAPBINWFRNESEIEHSHQLPQDDSGNYSLTIIDAYAEADSGEYKCAKNKIGK 825
QY 758 FPPGCKEAPLVSPFFLGGQAPPAKAPAPIDSKMGPQ-DISLQRPKPPGCSPPGS 816
DB 826 AHTVCCVRIEELS-----KRSKKIDGSKAPFRMQLP-TPEEV----- 864
QY 817 ASQASSSQVSSLRVGSSQVTEPGSPSLDAEGWTOEADLSDTPTLQRPQVQWTRKFSL 876
DB 865 -----QCADITLVCSVSGT-PHENIK---WTXDDKFDMSNKQVRHENGVCILH--II 911
QY 877 GORGAGVAGYGTFAFGDAGMLGQGPMMARITAMVSQSEEESEAEARASQSEEQE 936
DB 912 GAR-----DDDGRYCEAENIHGVAQSFSVVEIK-----EAYDKDHRV 950
QY 937 AAASPSLPQ-----VSAREPVVGRAPTRSSPEPTPEWED-----IQQVSLVQ 978
DB 951 PAFLEPLVNCSCGEMVLECCVTGKFIPI-----TWYKDGKLLIENRMLQ 999
QY 979 IRDLG-----DAAADTISLDISEVDPAYLNLSDLYDIKILPFFEMI 1021
DB 1000 YTRKGVSRNLNIMVNMNDGEBYTCEAVNSLGKDFTHCTVKVDMG-LSKTRLTPVR-- 1055
QY 1022 FRKVPKSAQP-----EPP---SPMAEELAE-----PPEPTWMPGELGP 1058
DB 1056 SRSRSRSPSVVSGEIQRPVTFELADATVTEGNRELEVEVDGPTPTPIEW---Y 1110
QY 1059 HAGLITRESEVDALLAEAVGRKWSRSL-FHFQGR--HLPLDEPAE----- 1108
DB 1111 HDG-----KLVAE-----SRLRTYFDGRVAFLLKIYEAHEHNGQYV 1147
QY 1109 -----LGLRER-----VKA--SVEHISRIILKGRPEGLEKEGPPRKKPLGLASFSLGKS 1155
DB 1148 CKVSNKLGAVETRAIVVEAPDAAEHVQM----- 1177
QY 1156 WPARTEFLRELSDEFVL---QSVTLACQVSAQAPAAQATMSKDGAFL-BSSSRVLISAT 1211
DB 1178 -----PTFVKKQD---VLKTAGETATFTCCSYAPAAQVWHLNKGKALQQTNSYKTRLF 1231
QY 1212 LKRFQLLTILVVAEDLVGYTCSVSNALGTVTTGVLR-----KABEP----- 1254
DB 1232 DNTATLVIENTDBLCGTYTAVANNQFVDVHTSAQLTISGSEAKIAASLPYFIELKP 1291
QY 1255 -----SSSPGPDIGEVYADGVLLVWKPVESYGPVTYVOCSEGGSWTT 1298
DB 1292 KINVEGATLSIQADLNGSPIPEV-----VWLKDNSELVESDRIQMKCDGVNYQL 1341
QY 1299 LASDIFDCCYLTSKLSRGTYTFTTACVSKAGMGPYSPSEQVLLGGPSHLASEESQ-- 1356
DB 1342 LVRDV-----GLEDEGTYTI-TAENEKGI-----QNTESVSTKSEKVEKKEKKV 1388
QY 1357 -----GRSAOPLPS-----TKTFAFOTQIORGRFVSVROC-----WEK-ASG 1392
DB 1389 EKKDGGKKKPGRPGLPFRSGASKTEQVTMAPDASEGPADSYEVERCPCDQREWVSCGT 1448
QY 1393 RALAAKIIYPHP-----KDKTAVLREYALKGLRHPHLAQLHAAYLSPPHLVLIL 1442
DB 1449 KSLLEIKGLTPNTEYIFRVAGKNKQGLGEMSEMTSLTKTASVQGAQPTISQSKI--- 1505
QY 1443 ELCSGPPELLPLAERASYSSESEVKDYLWQMLSATQYLN-----QHILH 1486
DB 1506 -----IANDDEFEIATVFSGTPTSPVKWYKENQIVPDEKIDVATSTSSILN 1554
QY 1487 LDLRSENMIITEYNLLKVVLDLGNASQUSQ 1515
DB 1555 LKSOBEN---GTFNCLNIENELGQASACQ 1580

RESULT 7

A88852
 Protein unc-22 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C:Accession: A88852
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
 A:Reference number: A75000; PMID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
 A:Accession: A88852
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-6831 <STO>
 A:Cross-references: GB:chr_IV; PIDN:CAA98081.1; PID:G3881830; GSPDB:GN00022
 C:Genetics:
 A:Gene: unc-22
 A:Map position: 4
 C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology;
 Query Match 5.7%; Score 498; DB 2; Length 6831;
 Best Local Similarity 20.4%; Pred. No. 1.8e-08;
 Matches 394; Conservative 256; Mismatches 708; Indels 572; Gaps 91;
 QY 51 PGPPSQV-TIEDVQQTGTAQFE-AITGGPQSVTWYKDSVOLVSTR-LSQQQEGT 107
 DB 4508 PGPIFDNNKQITVRAGETIIRIPYAGGPKPIIDLFNGSPFENERTVVDVNRG- 4566
 QY 108 TYSLVLRHVASK-DAGVYTCLAQNTGGQVLCKAELLVLGDNEDP-----SEKQSHRRK 160
 DB 4567 --BIVITTTGSKSDAGPKYKISATNKYKDKTKLVNVLDAKPGTPIRATDIQADAMT 4624
 QY 161 LHSFYEVKEIRGVFGFKVRVQHKNKILCAAKTI-----PLSRTRAQAY 207
 DB 4625 L-SWRPPKONGDAITNYVEKTPGGDWTVGHVGTTLVRNLDANTPYEFVRANQ 4693
 QY 208 -----RERDILAALSHPLVT-----GLDQFFTRKTLILI----- 237
 DB 4684 YGVGEPLTDDAIVAKNPFDTGAPQPEAVETSBEAITLQWTRPTSDGGAPIQYVIEK 4743
 QY 238 LELCSSP-----ELLRLRYKGVVT-----EAEVKVYIQQLVGLVHLS----- 277
 DB 4744 REVGSFTWKAAGNTLIDTKHRTVGLTPKTYEFVRVAY-NAAGGEYSVMSVITADNA 4802
 QY 278 -----HGVHLHDI-----KPSNLMVHPAREDIKIDFGPAQN-ITPAE-----LQFSQ 320
 DB 4803 PTRPKINMGMLTRDILAYAGERAKILVPPAASPAPKV-TFSKGENKISPTDPRVKVEYS- 4860
 QY 321 YGSPFVSPRIQQNPFSEASDIWAGVLSYLSLTCSPFAGESDRATLLNVLRGVSW 380
 DB 4861 ----DEFLATITIK-----SELTD-----GGLYFVELENSQ-----GSDSASIRLKVVD-----XPA 4903
 QY 381 SPWAHLSEDADKFIKATLQAPQAPRSPAAQCLSH-----PW-----FLKSMP 423
 DB 4904 SPOHIRVEDIAPDC--CTLYWMPSPSDGSPITNYIVKELDLRHSQKWKVSSFVRNLN 4961
 QY 424 ABEAHFINTKQLFLARSRWORSL--MSYKSIILVMR--SIPELLRGPV----- 468
 DB 4962 YTVGGLIKONRYFRV-RAETQVGVSEPCELADVVAKYQVEVFNQPEAPTVRKDSTWA 5020
 QY 469 -----DPSLGVARHL-----CRDTGSSSSSSSSNNELAPFAKSLPPSVTHSPIL-- 517
 DB 5021 ELEWDPDRDGGSKIIIGYQVYRDTSSGRWINAKMDLSEQCCHARVTGLRQNGEFEFFRIIAK 5080
 QY 518 HPRGFLRPSASLPEEABASERSTEAPAPASPEGAGPP--AAQCVPRHSVIRSLFVHQ 574
 DB 5081 NAAGFSKPS--PPSERCQKSRFGP-----PGPIHVGAKSIGRHNCTIT- 5123
 QY 575 AGSPSEHGALAPGSRHRHARRHLKGGYIAGALPGLREPMEH-----RYLEBEAAREE 629
 DB 5124 -----WMAP-----LEDGG--SKITGVNVEIREYGSTLWTVASDYNVREP 5161

QY 630 QATL--LAKAPSE-TAURLPASGTHLAPGHSHLSLEHDSPTTPRSSACGEAQLRPSAP 686
 DB 5162 EFTVDKLRFNDFYFRVVAINAAG-----KGIFSLP 5192
 QY 687 SGGAPIRDMGHGQKQLPSTGCHP-GTAQP-----ERPSDPSW---GQAP 730
 DB 5193 SGPIKIQESS--GSR--POIVVKEDTAQPNRAVETCEAVGPEPETARMLNRELPL 5247
 QY 731 FCHPKGSAFQGCSPHFAVACPFGSPGSCCK-----EAPLVSPSPFLQOP 778
 DB 5248 EBSRYFEASD-----GVYKFTKEVMDIDAGEYTVSVSNPY-GSD 5287
 QY 779 QAPPAPAKASPLDCKMGPGLDIPRPRKPGPCSPGSAQASQSSQVSSSLRVGSSQVTE 838
 DB 5288 TATANLVQAPVIEKDVNTI-LP-----SGDLVRLAIYES--GTA 5326
 QY 839 PGPSLDAEGWTQAEADLSDTPTLQRPQEQ-----VTMRKFSLGGRGY-----AGVAGYT 890
 DB 5327 P-----FRHSLVLRBEIDMDHRTIIRIVDFDHILITIPALSVREAGRYEYTVSNDGSEAT 5382
 QY 891 PAFGSDAGGM-----LQQGPMMARIAWA-----VSQSE--- 918
 DB 5383 TGFMLNVTGLPEAPOPLHISNIGPSTATLSWRPPTVTDGSKITSYVVEKRLDKDEWVT 5442
 QY 919 -----ESEOBEARAESQSESQOEARAESPLPOVSARVPVGVGAPTRSS 962
 DB 5443 VTSNVKDMNYIVTGLFENHEYEFVSAQNGIGA-----PLVSEHPI--IARLPDPP 5494
 QY 963 PPTFWEDI-----GVSLVQIARDLSG-----DAEAADTTLSLDISEYDPAYLN 1005
 DB 5495 TSPNLLEIVQVGDDVTLISWQRPISDGGRLRGYIVEXQEEHDEWFRCONQNPNNYN 5554
 QY 1006 LSDLVD--IKLVPF-----EFMLPRKPKSAQPEPPSMA----- 1038
 DB 5555 VFNLDIGRKRYRFAVNDAGLSLAEIDQTLFOASSGEGFKIVPLSLDLINEEVRGCVT 5614
 QY 1039 -EEELAEPEPTWPW--PGEL-----OPHAGLEITE-ESEVDLALLAAVAVR 1082
 DB 5615 FECEISGSPRPEYRPFKCKELVDTSKYLINKDKQVLIINDLTSDDADYTCRATN-- 5672
 QY 1083 KKKMSSPSRSLFPHGRHLPLDEPAELGLRERVASVEHISRLKGRPEGLEKGGPREKK 1142
 DB 5673 ----SSGTRS-----TRANLEIKTKPRVFI-----PKYH 5698
 QY 1143 PGLASFRLSGLSKWDRAPTFLRELSDETIVLQGVSTLACQVSAQAPAAQATWSKDGABLES 1202
 DB 5699 GGYEAQK-----GETIELKIPYKAYPQGEARWTKDGEKIEN 5734
 QY 1203 SSRVLISATLKNFQLTTLVVVAEDLVYTCVSNALGTVTTTGVLRKAERPPSSPCPD 1262
 DB 5735 NKSFSIT-TDDKFAFLRISNASREDYGEYRVVENSVDGSGTIVTVADVPEPRFPII 5793
 QY 1263 GEVYAGVLLVWKPVEYSGPVYIVQCSLE-----GGSWTTLASDIFDCCVLTSK-LSRG 1316
 DB 5794 ENILDEAVILSKPPALDGG-SLVNTYTIKREAMGGSWSPCAJSRY--TYTTTIEGLRAG 5850
 QY 1317 GYTTERTACVSKAGMPYSSPSEOVLLGGPS--HLASESESQGRSAQ----- 1361
 DB 5851 KYEYRIIAENKHGSKCEPTAPVLIPGDERKRRGYDVBQGIYRKGTVSSNYDNY 5910
 QY 1362 -----PLP-----STKTFATQIQGRFSVVRQCKEASGRALAAK-IIPYPH 1404
 DB 5911 VFDIMKQYPOPEVETKHDHLDHDIHEELGTGAFGVYHVRTERATGNFAAEVMTPE 5970
 QY 1405 KDKTAVLREYEAALKGRPHLAQLHAAYLSPRHLVILELCSGPPELLPCLA-ERASYSSES 1463
 DB 5971 SKETVRKEIQTMVSLVRHPTLVNLDHAFEDDNEWMVIEFMSGGELFEKVADEHNKMSD 6030
 QY 1464 EYKDYLMQWLSATQVLLHNOHILHLDLASENMIIT--FYNLLKVVLDLNAQSLSOEKVLPS 1521
 DB 6031 EAVETMROVCKGLCHMHENNYVHLDLKPENIMFTKSNELKLDLDFGLTAFLD-----PK 6085
 QY 1522 DRFK--DYLETMAPLELGGQGAVPQTDIWAIGVTAFLMISAEPVSESGARDLQRLRK 1578

DB 6086 QSVKVTGTAEAPPAEVAAGKPVGYVYTDWMSGVLSYIIULSGLSPPGGENDDETNRNKS 6145
QY 1579 GLVRLSRC-----YAGLSGGAVALFSTTCAQPWGRPCASSCLQCPLWTEGPACSR 1630
DDB 6146 -----CDWNMDSAFSGISDEGKDFIRKLLADENTMTIQALEHPWLT-PGNAPGR 6197
QY 1631 PAPVTFPTAR 1640
DDB 6198 DSQI--PSR 6205

RESULT 8
S57242
twitchin [similarity] - Caenorhabditis elegans
N:Alternate names: myosin-regulating protein
N:Contains: protein kinase (EC 2.7.1.-)
C:Species: Caenorhabditis elegans
C>Date: 28-Oct-1995 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S57242; S07571; S6797; S57218; T27934; T28030
R:Benian, G.M.; L'Hernault, S.W.; Morris, M.E.
A:Description: Additional sequence complexity within twitching of Caenorhabditis elegans
A:Reference number: S57242
A:Accession: S57242
A:Molecule type: DNA
A:Residues: 1-6839 <BEN1>
A:Cross-references: EMBL:L10351
A:Experimental source: var. Bristol
R:Benian, G.
A:Submitted to the EMBL Data Library, November 1989
A:Reference number: S07571
A:Accession: S07571
A:Molecule type: DNA
A:Residues: 792-6839 <BEN2>
A:Cross-references: EMBL:X15423; NID:g6897; PIDN:CAA33463.1; PID:g6898
A:Experimental source: var. Bristol
R:Benian, G.M.; Kiff, J.E.; Neckelmann, N.; Moerman, D.G.; Waterston, R.H.
Nature 342, 45-50, 1989
A:Title: Sequence of an unusually large protein implicated in regulation of myosin activation
A:Reference number: S06797; MUID:90044042; PMID:2812002
A:Accession: S06797
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 806-1175; 1178-1998, 'Y' 2000-3040, 'I' 3042-3335, 'I' 3337-5693; 5696-6359, 'I' 6359-6360
A:Cross-references: EMBL:X15423
A:Experimental source: var. Bristol
R:Benian, G.M.; L'Hernault, S.W.; Morris, M.E.
Genetics 134, 1097-1104, 1993
A:Title: Additional sequence complexity in the muscle gene, unc-22, and its encoded protein
A:Reference number: S57218; MUID:93387664; PMID:8397135
A:Accession: S57218
A:Molecule type: DNA
A:Residues: 2-99; 108-194, 'Q' 196-206; 374-468; 658-753 <BEN4>
A:Experimental source: var. Bristol
R:White, S.
A:Submitted to the EMBL Data Library, May 1996
A:Reference number: Z20442
A:Accession: T27934
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 'MGIFGKKCKQ', 19-6839 <W1>
A:Cross-references: EMBL:Z73897; PIDN:CAA98064.1; GSPDB:GN00022; CESP:ZK617.1a
A:Experimental source: clone ZK617
R:Haris, B.
A:Submitted to the EMBL Data Library, May 1996
A:Reference number: Z20458
A:Accession: T28030
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 'MGIFGKKCKQ', 19-6839 <W12>
A:Cross-references: EMBL:Z73899; PIDN:CAA98081.1; GSPDB:GN00022; CESP:ZK617.1a
A:Experimental source: clone ZK829

C-Comment: Lack of unc-22 leads to a constant twitching of the body muscles.

C-Genetics: C:unc-22; CESP:ZK617.1a
A:Gene: unc-22; CESP:ZK617.1a
A:Map position: 4
A:Introns: 18/3; 69/3; 143/2; 176/3; 264/2; 387/3; 413/2; 471/1; 516/3; 550/3; 582/3; 616/3; 669/1/3; 6776/1; 6808/3
C:Superfamily: twichin; fibronectin type III repeat homology; immunoglobulin homology;
C:Keywords: ATP, autophosphorylation; duplication; muscle; phosphotransferase; serine/ti
F:806-898-899-990-991-1093-1084-1175-1178-1273-1474-1587-1770-1864-2066-2158-2358-2450-;
96-5790-6263-6356-6386-6478-6541-6635-6649-6742-6743-6898-
F:1274-1372-1373-1473-1568-1670-1671-1769-1865-1964-1985-2065-2159-2258-2259-2357-2451-;
22-4215-4313-4314-4415-4416-4516-4612-4710-4711-4811-4908-5009-5010-5109-5110-5210-5399-
F:5940-6197/Domain: protein kinase homology <kin>
F:5948-5956/Region: region kinase ATP-binding motif
F:5971/Active site: Lys #status predicted

	Query Match	5.7%; Score 498; DB 2; Length 6839;
	Best Local Similarity	20.4%; Pred. No. 1.8e-08;
	Matches 394; Conservative 256; Mismatches 708; Indels 572; Gaps 91;	
QY	51	PGPPSMQV-TIEDVQAQTGGTAQFE-AIEGDPQPSVTWYKDSVLQVDPSTR-LSQQQEGT 107
DB	4516	PGRPFDINNKIDITVRAGETIQTIRPVAGGNPKPIIDLFNGNSPIFENERTVVDVNPQ- 4574
QY	108	TYSVLVRHVASK--DAGVYTCLAGNTGGQVLCKAELLVLGGDNEPD-----SEKSHRRK 160
DB	4575	--EIVTTTGGKSDAGYKISATNKYGRQTKLVNFVLDAPKGTGPIRATIDQAMT 4632
QY	161	LHSFYEVKEIGRGVFGFKRVQHKGNKILCAAKPI-----PLESRTRAQAY 207
DB	4633	L-SWRPPKONGGDALTNNVWEKRTPGGDWTVGHPVGTTLVRNLNDANTPYEFVRAENQ 4691
QY	208	-----REEDILAALSHPLVNT-----GLLDDQETKTKTLILI----- 237
DB	4692	YGVGEPLTDDAIVAKNFDTPGAPQGEAVETSEATLTOWRFTSDGGAPIQGVVTEK 4751
QY	238	LELCSS-----ELDLRYLRYGVVT-----EAEVKYIIOQLVEGLHYLHS----- 277
DB	4752	REVSGSTETKAAFGNILLDTKRVGTGLTPKTYEYFRAAY-NAAQGEYSVNSVPITADNA 4810
QY	278	-----HGVLLHDI-----KPSNLMVHPAREDIKICDFGAQN-ITPAE-----LQFSQ 320
DB	4811	PTRPKINMGMLTRDILAYAGSRKILVPFASPAKPV-TFSKGENKISPTDPRVKVEYS- 4868
QY	321	YGSPEFVSPETIQNPVSEASDIWAMGVISYLSITCSPPFAGESDRATLLNVLEGRVWS 380
DB	4869	----DFLATLTIK-----SELT-----GGLYFEVLENSQ-----GSDSASIRLKVVD---KPA 4911
QY	381	SPMAAHLSEDAKDFIKATLQAPARPASAAQCLSH-----PW-----FLKSNP 423
DB	4912	SPQIRVEDIAPDC--CTLYWNPSSDGGSPITNVIYEKLDLRHSDGKWKVSSPVRLN 4969
QY	424	AEBAHFINTKQLKELLARSQWRSI-----MSYKSIILWVR---SIPELLRGPP----- 468
DB	4970	YTVGGLIKONRYFRFV-RAETQYGVSEPCELADVVAKYQFVFNQPEAPTVRDKDSTWA 5028
QY	469	----DSPSLGVARHL-----CRDTGGSSSSSSSSDNLAPARAKSLPPSPVTHSPLL-- 517
DB	5029	ELEWDPDPDGGSKLIGYQVQYRDTSSGRWINAKWDLSEQCHARVTGLRQNGEFEPRIIAK 5088
QY	518	HPRGFLPASPASLPEAEASERSTAPAPSPGAGPP---AAQCVPRHVSIRSLFVHQ 574
DB	5089	NAAGFSPKPS---PPSERQLKSRFGP-----PGPEIHVGAKSIGRNHCTIT----- 5131
QY	575	AGEGPEHGALAPGRRHPRARRHLLKGGYIAGALPGLREPLMEH-----RVLESEAAREE 629
DB	5132	-----WMAP-----LEDGG---SKITGYNVEIREYGSTLTWTVASDYNVREP 5169
QY	630	QATL--LAKASFE-TALRLPASGTHLAPGHSLSLEHDSPTPRPSSACGEAQLPSAP 686
DB	5170	EFTVDKLRFFNDYFRRVVAINAAG-----KGISLP 5200
QY	687	SGGAPIRDMGHFGQSKLPSTGCHHP-GTAQSP-----BRPSPDSPW---GQAP 730

Db 5201 SGPTKIQESG---GSR--PQIVVKPETAQPNYRAVFTCEAVGREPTARWLRNGREL 5255
Qy 731 FCHPKQSAPOEGCSGPHPAVACPSPGPPGSCK-----EAPLVSPSPFLGP 778
Db 5256 ESSRYRFEASD-----GVYKFTIKVMDIDAGEYTVVEVSNPY-GSD 5295
Qy 779 QAPAPAKASPLDSKMGCGDISLPGRPKPCPGSCPSGASQSSQVSSLRVSSQVGT 838
Db 5296 TATANLVVQAPFVIEKDVNTI-LP-----SGDLVRLKIYFS--GTA 5334
Qy 839 PGPSLDAEGMTQEAEDLSSTPTLQRFQEQ---VTNKRFSLGRGQY----AGVAGYGT 890
Db 5335 P-----FRHSLVLRNEEDMDHPTIRIVEFDDHILITIPALSVREAGRYEYTVSNDSEAT 5390
Qy 891 FAFGSDAGM-----LGQCPMKWATAWA-----VSQSR--- 918
Db 5391 TGFNLVNTGLPEAPQGLPHISNIGPSTATLSWRPPVTDGSKITSYVVERKDLKDEWVT 5450
Qy 919 -----BEEQBEARAEQSEQBEARAESELPQVSARPPVEVGRAPTRSS 962
Db 5451 VTSNVKQWYIVTGLFENHGEYFRVSAQENIGCA-----PLVSEHPI--IARLPDP 5502
Qy 963 PEPTPWEDI---GOVSLVQIRDLG-----DAAADTISLDISEVDPAVLN 1005
Db 5503 TSPNLNLEIVQVGGDYVTLVSWORPLSDGGRGRLRGYIVKEQBEEHDEWFRQNPSPNNYN 5562
Qy 1006 LSDLYD---IKYLP-----EFMIFRKVPKSAQPEPSPMA----- 1038
Db 5563 VPNIIDGRKYRYFVAVNDAGLSDLAELDQTLFQASGEGEKIVSPISDLNEBVGRCVT 5622
Qy 1039 -BEEIAEPPEPTWPH---PGBL-----GPHAGLEITE-ESDVDALLAEAAVGR 1082
Db 5623 FCEITSGSPREYRWFKCKELVDTSKYTLNKGDKQVLIINDLTSDDADBYTCRATN-- 5680
Qy 1083 KRKWSPPSRLFFHPPGRHLPLDEPAELGRVRKASVEHISRLKPRGLEKEGPPRKK 1142
Db 5681 ----SSGTRS-----TRANLEIKTPRVFI-----PKYH 5706
Qy 1143 PGLASFRSLGKSWDRAPTFIRELSDETIVLQGSVTLACQVSAQAPAAQWSDGAPLES 1202
Db 5707 GGYEAQK-----GETIELKIPYKAYPOGEARWTKDGKIKEN 5742
Qy 1203 SSRVLISATLKNFOLLITLVVAEDLVYTCVSNALGTVTITGLRKAERPSPPCPDI 1262
Db 5743 NSKFSIT-TDDKFAFLRISNAREDYGEYRVVENSVDGSGTVNVTADVPEPRFPII 5801
Qy 1263 GEVYADGVLLVWKPVESGPTVIYVQCSLE-----GGSWTTLASDIFDCCYLTSK-LSRG 1316
Db 5802 ENILDEAVILSNKPPALDGG-SLVNTVYTIKKEAMGGGWSPCAKSRY--TYTTIEGLRAG 5858
Qy 1317 GYTTTRTACVSKAGMPYSSPSQVLLGSPS---HLASEESQGRSAQ----- 1361
Db 5859 KQYEFRIIAENKHGQSKCEPTAPVLIIPGDERKRRGYDVEQGIKVRGKGTVSSNDYNY 5918
Qy 1362 -----PLP-----STKTFATQIQGRFVSVRQWKAAGRALAAK-IIPYHP 1404
Db 5919 VFDIMKQYYPQVEIKHDHVLHDYDIHEELGTGAFGVVHRVTERATGNFAAKFVMTPE 5978
Qy 1405 KDKTAVLREYALKGLRPHPLAQLAAVLSPRHVLILELCSGPPELLPCLA-ERASYS 1463
Db 5979 SDKETVRKEIQTMSVLRHPTLVNLHDAFEDDNDNMVYIYEFMSGGELFEKVADENKMSD 6038
Qy 1464 EVKDYLMQMLSATQYLHNOHILHDLRSNMIIIT-EVNLKAVLDLGNAAQSLSQEKVLPS 1521
Db 6039 EAVEYWRQCKGLCHGHNENYVHLDLKPENIMFTTKRSNELKLIDFGLTAHLD-----PK 6093
Qy 1522 DKFK-----DYLETWAPELLGQGAVFQTDIWAIGVTAFIMLSAEYVPSSEGARDLQGRUK 1578
Db 6094 QSVKVTGTGAEPAAPEVAEGKPGVYVTDWMSVGLSYILLGSLSPFGGENDDETUNYKS 6153
Qy 1579 GLVRLSRC-----YAGLSGGAFLRSTLCAQPWGRPCASSCLQCPWLTTEGPACSR 1630

Db 6154 -----CDWNMDSDAFSGISEDGKDFIRKULLADPNTRMTHTHQALEHPWLT-PGNAPGR 6205
Qy 1631 PAPVTFTPTAR 1640
Db 6206 DSQI--PSSR 6213
RESULT 9
T27935
hypothetical protein ZK617.1b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
C:Accession: T27935; T28031
R:White, S.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z20442
A:Accession: T27935
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7160 <W12>
A:Cross-references: EMBL:Z73997; PIDN:CAA98065.1; GSPDB:GN00022; CESP:ZK617.1b
A:Experimental source: clone ZK617
R:Harris, B.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z20458
A:Accession: T28031
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7160 <W12>
A:Cross-references: EMBL:Z73999; PIDN:CAA98082.1; GSPDB:GN00022; CESP:ZK617.1b
A:Experimental source: clone ZK829
C:Genetics:
A:Gene: CESP:ZK617.1b
A:Map position: 4
A:Introns: 10/3; 61/3; 135/2; 168/3; 256/2; 379/3; 405/2; 463/1; 508/3; 542/3; 574/3; 591/3
3067/1; 3141/3; 3269/1; 6473/3; 7012/3; 7037/1; 7129/3
C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology; I
Query Match 5.7%; Score 498; DB 2; Length 7160;
Best Local Similarity 20.4%; Pred. No. 1.9e-08;
Matches 394; Conservative 256; Mismatches 708; Indels 572; Gaps 91;
Qy 51 PGPPSQV-TIEDVQATGCTAQFE-AIEGPPQPSVTWYKDSVQLVDSTR-LSQQOEGT 107
Db 4837 PGRIPIFINLKDITVRAGETIQRIPYAGNPKPIIDLFNGNSPIFENERTVVDVNP- 4895
Qy 108 TYSVLVLRHASK--DAGVVTCLAQNTGGVLCVCAELLVLDGNEPD-----SEKQSHRRK 160
Db 4896 --EIVITTTGSKRSDAGPKISATNKYKDTCKLVFVLDAPKPTGPIRATDIQADMT 4953
Qy 161 LHSFYEVKBEIGRGVFGVKRVQHKGNKILCAAKFI-----PLRSTRQAQY 207
Db 4954 L-SWRPPKONGGDAITNYVVEKRTPGDWVTGHPVGTTLRVNLDANTPYEFVRAENQ 5012
Qy 208 -----RERDILAALSHPLVT-----GLDQFETKTLILI----- 237
Db 5013 YGVGEPLTDDAIVAKNPFDPGAPQPEAVETSEAILOWTRPTSDGAPIQGVIEK 5072
Qy 238 LELCSSS-----ELLRLYRKGVVT-----EAEVKYIQQVLEGLHYLHS----- 277
Db 5073 REVGSTEWTKAAGNILDTHKRVTLGLTPKTYEFRVAAY-NAAGQGEYSVNSVPIITADNA 5131
Qy 278 -----HGVHLIDI-----KPSNLMVHPAREDIKICDGPAGN-ITPAE-----LQFSQ 320
Db 5132 PTRFKINMGLTRDILAYAGERAKILVFFAASAPAKY-TFSKGENKISITDPRAKVEYS- 5189
Qy 321 YGSPFVSPRIIOQNVPVSEASDIWANGVTSYLSLTCSPPFAGSDRATLNLVLEGRVWS 380
Db 5190 ----DPLATLTLEK-----SELT-----GGLYFVELENSQ-----GSDSASIRLKWVD--KPA 5232
Qy 381 SPMAHLSDEAKDFIKATLQRAQAPASAAQCLSH-----PW-----FLKSM 423
Db 5233 SPQHIRVEDIAPDC--CTLYMPPSSDGGSPITNYIVEKILDLRHSQKWEKVSFVRNLN 5290


```

Db 212 E---SKEAGEKAPQADQAKVQDTSRGIFQAVPSE-----RPRPEVGOALCLPAR 261
QY 138 KAEILLVGGDNEPDKQSHR-----RKLHSFYEV--KEEIGRVFGFVKRVQHGKNGKI 189
Db 262 EEDCFQILDCCPPAPPPHRIVELRTGNTSSBSMSKALGGKFGAGVCTCTEKSTGL 321
QY 190 LCAAKFTPLRS-RTRAQAYRERDILALSHPLVTGLDQFETKTLILILELCSSELLD 248
Db 322 KLAARVKKQTPKQKEMVMEIEVMQNLNHNLIQLYAAIETPHEIYLFMEYIEGGELFE 381
QY 249 RLYRKGV-VTEAEVKVYIQQLVEGLHYLHSHGVHLHDKPSNLMVHPAREDIKCDFFG 307
Db 382 RIVDEYHLETVDNWVFRQICDGLFHKVRVHLHDKPENILCVNTTGHVYKLIIDFGL 441
QY 308 AONTTPAELQPSQVSGSPFVSPETIQNQPVSEADSIWAMGVISYLSITCSPSPAGESDRA 367
Db 442 ARRYNPNEKLVNFTGPELSPVNVYDQISDKTDMWSLGVITYMLLSGLSPFLGDDDE 501
QY 368 TLINVLGVRGWSVPMALHSEDAKFIKATLQAPQAPQAPSAACLSHPWFLKSPAEAA 427
Db 502 TLINVLGNTWTFDEETFAVSDAEKDFVSNLIVKEQGRVSAACLSHPW-LNKL-AEKA 559
QY 428 HFINTK-----QLKFLARSRWQSLMS 450
Db 560 KCNRRRLKSQLLLKYLKMRWKKNFIA 587

RESULT 11
S49128
A:Description: Autophosphorylation of molluscan twitchin and interaction of its kinase d
A:Reference number: S49128
A:Accession: S49128
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-451 <HE1>
A:Cross-references: EMBL:Z30161; NID:G509412; PIDN:CAA82911.1; PID:G509413
R:Heierhorst, J.; Probst, W.C.; Vilim, F.S.; Buku, A.; Weiss, K.R.
J. Biol. Chem. 269, 21086-21093, 1994
A:Title: Autophosphorylation of molluscan twitchin and interaction of its kinase domain
A:Reference number: A53923; MUID:94342273; PMID:8063728
A:Accession: A53923
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 5-372 <HE2>
A:Cross-references: GB:Z30161
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F:45-302/Domain: protein kinase homology <KIN>
F:53-61/Region: protein kinase ATP-binding motif

Query Match 5.3%; Score 463.5; DB 2; Length 451;
Best Local Similarity 38.3%; Pred. No. 1.9e-08;
Matches 105; Conservative 43; Mismatches 115; Indels 11; Gaps 4;

QY 149 EPDSKQSHRKLHSFYEVKEIGRVFGFVKRVQHGKNGKILCAAKFTI---PLRSRTFA 204
Db 34 QPVEVKGQ---SCDYDIDLELGGAGFVGVHRCVEKATGRVFEAKFTNPPLDKYTL-- 88
QY 205 QAYRERDILALSHPLVTGLDQFETKTLILILELCSSELLDRLYRKGV-VTEARVYV 263
Db 89 -VKNVEISINNQLHHPKLIHLDAFEDKYDMLILFLSGGELFDRIAAEDYKMSAEVYN 147
QY 264 YIQQLVEGLHYLHSHGVHLHDKPSNLMVHPAREDIKCDGFAQNTTPAELQPSQVGS 323
Db 148 YMRQACEGLKHHHEHIVHLDIKPENIMCETKASVSKIIDFGLATKUNPDDIVKVTAT 207
QY 324 PFVSPETIQNQPVSEADSIWAMGVISYLSITCSPSPAGESDRA TLINVLGVRGWSVPM 383

```

```

Db 208 AFAFAPEIVDRPEVGFYTDMAIGVIGVLGSLSPFFAGEDDLETQNVRCDNFEDEDA 267
QY 384 AAHLSEDAKFIKATLQAPQAPQAPSAACLSHPW 417
Db 268 FSSVSPKADFIKLLQKEPRKRLTVHDALEHPW 301

RESULT 12
I37275
A:Description: death-associated protein kinase (EC 2.7.1.1) - human
N:Alternate names: calmodulin-dependent protein kinase homology; DAP kinase
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 10-Jul-1998
C:Accession: I37275; S39269
R:Beiss, L.P.; Feinstein, E.; Berissi, H.; Cohen, O.; Kimchi, A.
Genes Dev. 9, 15-30, 1995
A:Title: Identification of a novel serine/threonine kinase and a novel 15-kD protein as ]
A:Reference number: A55614; MUID:95129831; PMID:7828849
A:Accession: I37275
A:Molecule type: mRNA
A:Residues: 1-1423 <RES>
A:Cross-references: EMBL:X76104; NID:G434846; PID:G434847
C:Genetics:
A:Gene: GDB:DAPK1; DAPK
A:Cross-references: GDB:555932; OMIM:600831
A:Map position: 9q34.1-9q34.1
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A:Pathway: apoptosis
C:Note: activity is calmodulin dependent
C:Superfamily: death-associated protein kinase; ankyrin repeat homology; protein kinase I
C:Keywords: apoptosis; ATP; calmodulin binding; phosphotransferase; serine/threonine-spe
F:11-267/Domain: protein kinase homology <KIN>
F:19-27/Region: protein kinase ATP-binding motif
F:285-308/Region: calmodulin binding #status Predicted
F:370-402/Domain: ankyrin repeat homology <AN1>
F:403-435/Domain: ankyrin repeat homology <AN2>
F:436-468/Domain: ankyrin repeat homology <AN3>
F:470-502/Domain: ankyrin repeat homology <AN4>
F:503-535/Domain: ankyrin repeat homology <AN5>
F:536-568/Domain: ankyrin repeat homology <AN6>
F:569-601/Domain: ankyrin repeat homology <AN7>
F:602-634/Domain: ankyrin repeat homology <AN8>
F:42,64,139,141/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 5.2%; Score 456; DB 1; Length 1423;
Best Local Similarity 33.1%; Pred. No. 9.3e-08;
Matches 109; Conservative 65; Mismatches 119; Indels 36; Gaps 7;

QY 158 RKLSHFYEVKEIGRVFGFVKRVQHGKNGKILCAAKFTPLR---SR---TRAQAYR 210
Db 6 QENVDDYDTGEEELSGSQFAVVKCKREKSTGTQYPAKFIKRTKSRRGVSEEDIEREV 65
QY 211 DILAALSHPLVTGLDQFETKTLILILELCSSELLDRLYRKGVVTAEVKVYIQQLVE 270
Db 66 SILKEICHNPVITLHEVYENKTDVILILELVAGGELFDLAEKESLTEBEATEFLKQILN 125
QY 271 GLHYLHSHGVHLHDKPSNLMV--HPAREDIKCDGFA-QNITPAELQFSQVSGPEFV 327
Db 126 GVYLYLSLQTAHFDLPENIMLLDRNVKPRIKIIDFGNEFKNI-----FGPEFV 176
QY 328 SPETIQNQPVSEADSIWAMGVISYLSITCSPSPAGESDRA TLINVLGVRGWSVPMALH 387
Db 177 APVINYEPFLGLADWMSIGVITYILLSGASPPLGDTKQETLANVSAVNYEDEFVSNT 236
QY 388 SEDAKPIKATLQAPQAPQAPSAACLSHPWF----LKSPAEAAHFTINTKQLKFLARSR 443
Db 237 SALAKDFIRLLVKDPKRRMTIQDSLQHPWIKPDKTQQAALSRKASAVNMEKFKKFAARK 296
QY 444 WQ-----RSLMSYKSLVMS 459
Db 297 WKQSVRLISLCQLRSRSLSRSNMVAR 325

```

RESULT 13

T13741
 hypophyseal protein 22S.8 - fruit fly (*Drosophila melanogaster*)
 CSpecies: *Drosophila melanogaster*
 CDate: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
 CAccession: T13741
 C:Submitted to the EMBL Data Library, April 1999
 A:Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.
 A:Reference number: Z17668
 A:Accession: T13741
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1398 <MUR>
 A:Cross-references: EMBL:AL031765; NID:e1371523; PID:e1329905; PIDN:CAA21125.1
 C:Genetics:
 A:Cross-references: FlyBase:FBgn0000667
 A:Introns: 205/3; 227/1; 322/3; 688/3; 782/3; 814/2; 1363/3
 A:Note: EG:22E5.8

Query Match	5.2%	Score 454.5;	DB 2;	Length 1398;
Best Local Similarity	22.4%;	Pred. No. 1e-07;		
Matches 262;	Conservative 149;	Mismatches 467;	Indels 223;	Gaps 47;

[illegible]

5.2%; Score 451; DB 2; Length 371;

Best Local Similarity	35.2%	Pred. No. 4e-08;
Matches 106; Conservative	68; Mismatches 111;	Indels 16; Gaps 7;

```

162 HSFYEV--KEEIGRGVFGVKRVQHKGNKTLCAAKFIPILRSR---TCAQAYPERDIL--AAL 216
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
29  NNFFYTLTPKEUGRKFAVVQQCTSKSTGQGYAAKFKKRRGGQDCQRABIIHEHTAVELAR 88
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
217 SHPIVTGLLDGFGRKTLIILILELCSSELLDLRLYRK--GVVTRAEVKVVIQQLEVLHY 274
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
89  SCFPHVINLHVEYTATEIILLVLEYAGGSGIFNLCLPELAEMVSNDVRIRIKQLLEGVHY 148
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
275 LHSHGVLHLDTKPSNTILM--VHPAREDKICDFGAQNITPAELQFSQYGSPFVSPSEII 332
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
149- LHQNNIVHLDDKPONILLSSIYPL-GDIKIIVFGMSRKIGNASLERIMGTPTSYATAEII 207
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 14

UC7733
death associated protein kinase related apoptosis inducing protein kinase 2 homolog, DR1
C.Species: Rattus norvegicus (Norway rat)
C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 14-Dec-2001
C.Accession: UC7733; PC7178
C.Author: Matsumoto, M.; Miyake, Y.; Nagita, M.; Inoue, H.; Shitakubo, D.; Takemoto, K.; Ohtsuka
J. Biochem. 130, 217-225; 2001
A.Title: A serine/threonine kinase which causes apoptosis-like cell death interacts with
A.Reference number: UC7733; MUID:21374135; PMID:11481038
A.Contents: Brain
A.Accession: UC7733
A.Molecule type: mRNA
A.Residues: 1-371 <Mat>
A.Accession: PC7178
A.Molecule type: protein
A.Residues: 227-371 <MA2>
C.Comment: This kinase, a novel calcineurin homologous protein (CHP) binding protein wit
involved in the transfer of Na+/H+ exchanger 1 process and in some apoptotic events.
C.Genetics:
C.Gene: drak2
C.Keywords: apoptosis

```

Db          : || | : ||| : :
567 RLKSQILLKCYLMKRRWKCNFIA 589

Search completed: April 23, 2004, 15:11:30
Job time : 41 secs

```

A: myosin-light-chain kinase (EC 2.7.1.117), skeletal muscle - rat
 C: Species: Rattus norvegicus (Norway rat)
 C: Date: 12-Jan-1989 #sequence_revision 12-Jan-1989 #text_change 19-Dec-1997
 C: Accession: A28798; A60441
 R: Roush, C.L.; Kennelly, P.J.; Giaccum, M.B.; Helleman, D.M.; Scott, J.D.; Krebs, E.G.
 J. Biol. Chem. 263, 10510-10516, 1988
 A: Title: Isolation of the cDNA encoding rat skeletal muscle myosin light chain kinase.
 A: Reference number: A28798; MUID:88273159; PMID:2839493
 A: Accession: A28798
 A: Molecule type: mRNA
 A: Residues: 1-610 <ROU>
 A: Cross-references: GB:J03986
 R: Herring, B.P.; Numally, M.H.; Gallagher, P.J.; Stull, J.T.
 Am. J. Physiol. 256, C399-C404, 1989
 A: Title: Molecular characterization of rat skeletal muscle myosin light chain kinase.
 A: Reference number: A60441; MUID:89148435; PMID:2465691
 A: Accession: A60441
 A: Status: not compared with conceptual translation
 A: Molecule type: mRNA
 A: Residues: 296-610 <HER>
 C: Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
 C: Keywords: ATP; calmodulin binding; muscle; phosphotransferase; serine/threonine-specif
 P: 297-554/Domain: protein kinase homology <KIN>
 P: 305-313/Region: protein kinase ATP-binding motif

Query Match	5.1%	Score	450;	DB	2;	Length	610;	
Best Local Similarity	28.9%;	Pred.	No. 6.7e-08;					
Matches	128;	Conservative	85;	Mismatches	178;	Indels	52;	Gaps
QY	43	LHS-SLPALPGPPSMQVITIEDVQAQTGTAQFAEAIIEG-----DPQPS-----VTWY	88					
Db	164	LHSPSCPAIISCSEKTLAMKPLSETT-----ELIFAGVSETPPDPGPAGKDEGGNTL	217					
QY	89	KDSVOLVDSTRLSQO--QEGTTYSLVLRHVAKDAQVYTCCLAQNTGQVLC-----XAELL	142					
Db	218	ADGKEEAEGAOEAKQVQDTSORIGFQAVPSEAEV-----GOALCLTAKESDCF	268					
QY	143	VILGGNEEDPSEKQSHR-----RKLSHSEYV--KEEIGRVGFGVFKRVHKGKILCAAK	194					
Db	269	QILDDCPFPAPFPFHRIIVELRTGNVSSEFSMNSKEALGGGKFGAVCTCTERSTGLKLAAK	328					
QY	195	FIPLS-RTRAQAVREROIILAALSHPLVTGLDQFETKTLVILILECSSEELLDRLYRK	253					
Db	329	VIKQTQPKDKEMVLIEIVMQLNHRNLIQLYSATIESHEILFMVEYIEGGELPERIVDE	389					
QY	254	GV-VTEAEKVYIQLVSLGHVLHSHGVHLIDKPSNILMVHPAREDIKICDFFAQNIT	312					
Db	389	DYQLTEVDTMFVRQICDGIIFPMHKRVHLHDLKPENLICVNTTCHLVKIIFDGLARRYN	448					
QY	313	PAELOFYSGSPFVSPRIIOQNPVSEASDIWAMGVISYLSLTCSPPAGESDRATLNV	372					
Db	449	PNEKLWNPGTDEFLSPFVWYDQISDXTDMWSLGVITYMLLSGLSPFLGDDDTETLNV	508					
QY	373	LEGKRVSWSPMAHLSDEADKOFIKATLQAPQAPSPAQCILSHPMFLKSPAEBAHFINT	432					
Db	509	LSANWYFBETFEAVSEDAKQFVSNLTIDQSQARMFAEQCLAHFW-LNNL-AEKAKCNR	566					
QY	433	K-----QLKFLIARSRWQRSLMS	450					

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2004, 14:55:57 ; Search time 24 Seconds
(without alignment)

3612.367 Million cell updates/sec

Title: US-10-697-263-2

Perfect score: 8740

Sequence: 1 MGCCCLGGCGSVAHVSQG.....RNREKRALLYKRNHLAQVR 1665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	592.5	6.8	3038	1	TRIO HUMAN
2	550.5	6.3	1176	1	KMLS BOVIN
3	544.5	6.2	1147	1	KMLS RABIT
4	541.5	6.2	1914	1	KMLS HUMAN
5	532.5	6.1	1906	1	KMLS CHICK
6	489.5	5.6	448	1	DAK3 MOUSE
7	489.5	5.6	448	1	DAK3 RAT
8	475.5	5.4	370	1	DAK1 MOUSE
9	471.5	5.4	1432	1	DAK1 HUMAN
10	469.5	5.4	1442	1	DAK1 MOUSE
11	468.5	5.4	593	1	KML2 HUMAN
12	466.5	5.3	454	1	DAK2 HUMAN
13	465.5	5.3	370	1	DAK2 RABIT
14	465.5	5.3	607	1	KML2 RABIT
15	463.5	5.3	1050	1	ULK1 HUMAN
16	460.5	5.3	438	1	KMLS SHEEP
17	451.5	5.2	371	1	S17B RAT
18	450.5	5.1	609	1	KML2 RAT
19	449.5	5.1	1051	1	ULK1 MOUSE
20	447.5	5.1	397	1	S17A RABIT
21	445.5	5.1	414	1	S17A HUMAN
22	441.5	5.1	374	1	KCCI RAT
23	440.5	5.0	424	1	KPSH HUMAN
24	438.5	5.0	372	1	S17B MOUSE
25	437.5	5.0	374	1	KCCI MOUSE
26	431.5	4.9	794	1	KILL HUMAN
27	428.5	4.9	372	1	S17B HUMAN
28	427.5	4.9	370	1	KCCI HUMAN
29	418.5	4.8	414	1	KCCI EMENI
30	413.5	4.7	779	1	SNIL MOUSE
31	412.5	4.7	752	1	MRKA HUMAN
32	406.5	4.6	602	1	CRK DAUCA
33	405.5	4.6	776	1	SNIL RAT

34 404.5 4.6 469 1 KCC4 MOUSE
35 404.5 4.6 786 1 SNIL HUMAN
36 399.5 4.6 473 1 KCC4 HUMAN
37 390.5 4.5 664 1 KCCB HUMAN
38 389.5 4.5 474 1 KCC4 RAT
39 389.5 4.5 240 1 KML2 MOUSE
40 387.5 4.4 733 1 K6A2 MOUSE
41 387.5 4.4 740 1 K6A3 HUMAN
42 387.5 4.4 740 1 K6A3 MOUSE
43 385.5 4.4 752 1 K6A1 CHICK
44 384.5 4.4 774 1 MRK2 MOUSE
45 384.5 4.4 735 1 K6A1 RAT

ALIGNMENTS

RESULT 1
TRIO HUMAN
ID TRIO HUMAN STANDARD; PRT; 3038 AA.
AC 075962; Q13458;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Triple functional domain protein (PTPRF interacting protein).
GN TRIO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
EN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fibroblast;
RX MEDLINE=96224308; PubMed=8643598;
RA Debant A., Serra-Pages C., Seipel K., O'Brien S., Tang M., Park S.-H.,
RA Streuli M.;
RT "The multidomain protein Trio binds the LAR transmembrane tyrosine
RT phosphatase, contains a protein kinase domain, and has separate rac-
RT specific and rho-specific guanine nucleotide exchange factor
RT domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:5466-5471(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Streuli M.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP STRUCTURE BY NMR OF 1227-1407 AND MUTAGENESIS.
RX MEDLINE=99003194; PubMed=9790533;
RA Liu X., Wang H., Eberstadt M., Schnuchel A., Olejniczak E.T.,
RA Meadows R.P., Schkeryantz J.M., Janowick D.A., Harlan J.E.,
RA Harris E.A.S., Staunton D.E., Fesik S.W.;
RT "NMR structure and mutagenesis of the N-terminal Dbl homology domain
RT of the nucleotide exchange factor Trio.";
RL Cell 95:269-277(1998).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99276567; PubMed=10341202;
RA Seipel K., Medley Q.G., Kedesha N.L., Zhang X.A., O'Brien S.P.,
RA Serra-Pages C., Hemler M.E., Streuli M.;
RT "Trio amino-terminal guanine nucleotide exchange factor domain
RT expression promotes actin cytoskeleton reorganization, cell migration
RT and anchorage-independent cell growth.";
RL J. Cell Sci. 112:1825-1834(1999).
CC -!- FUNCTION: PROMOTES THE EXCHANGE OF GDP BY GTP. TOGETHER WITH
CC LEUCOCYTE ANTIGEN-RELATED (LAR) PROTEIN, IT COULD PLAY A ROLE IN
CC COORDINATING CELL-MATRIX AND CYTOSKELETAL REARRANGEMENTS NECESSARY
CC FOR CELL MIGRATION AND CELL GROWTH.
CC -!- SUBUNIT: INTERACTS TO FORM A COMPLEX WITH LEUCOCYTE ANTIGEN
CC RELATED PROTEIN.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=075962-1; Sequence=Displayed;

Name=2;
 -!- isoId=075962-2; Sequence=VSP_004467, VSP_004468;
 -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,
 BRAIN, PANCREAS, PLACENTA, LIVER, KIDNEY AND LUNG.
 -!- DOMAIN: THE N-TERMINAL DBL/GEF DOMAIN SPECIFICALLY CATALYZES
 NUCLEOTIDE EXCHANGE FOR RAC1, LEADING TO THE ACTIVATION OF JUN
 KINASE AND THE PRODUCTION OF MEMBRANE RUFFLES. THE SECOND DBL/GEF
 DOMAIN IS AN EXCHANGE FACTOR FOR RHOA AND INDUCES THE FORMATION OF
 STRESS FIBERS.
 -!- PTM: Phosphorylated on serine residue(s).
 -!- SIMILARITY: Contains 1 CRAL-TRIO domain.
 -!- SIMILARITY: Contains 2 DBL-homology (DH) domains.
 -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 -!- SIMILARITY: Contains 2 PH domains.
 -!- SIMILARITY: Contains 1 SH3 domain.
 -!- SIMILARITY: Contains 4 spectrin repeats.
 -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SER/THR
 FAMILY OF PROTEIN KINASES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AF091395; AAC3042.1; -
 EMBL; U42390; AAC34245.1; -
 HSSP; Q63450; IA06.
 Genew; HGNC:12303; TRIO.
 MIM; 601993; -
 GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; TAS.
 GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
 GO; GO:0007185; P:transmembrane receptor protein tyrosine pho. .; TAS.
 InterPro; IPR001251; CRAL_TRIO_C.
 InterPro; IPR001331; GDS_CDC24.
 InterPro; IPR007110; Ig-like.
 InterPro; IPR003598; Ig_c2.
 InterPro; IPR001849; PH.
 InterPro; IPR000719; Prot_kinase.
 InterPro; IPR000213; RhoGEF.
 InterPro; IPR008271; Ser_thr_pkinase.
 InterPro; IPR002290; Ser_thr_pkinase.
 InterPro; IPR001452; SH3.
 InterPro; IPR002017; Spectrin.
 Pfam; PF00047; Ig; 1.
 Pfam; PF000159; PH; 2.
 Pfam; PF00069; pkinase; 1.
 Pfam; PF00621; RhoGEF; 2.
 Pfam; PF00019; SH3; 1.
 Pfam; PF00435; spectrin; 6.
 ProDom; PD000001; Prot_kinase; 1.
 SMART; SM00408; IGC2; 1.
 SMART; SM00233; PH; 2.
 SMART; SM00325; RhoGEF; 2.
 SMART; SM00220; S_TKC; 1.
 SMART; SMC0516; SEC14; 1.
 SMART; SMC0326; SH3; 2.
 SMART; SMC0150; SPEC; 6.
 PROSITE; PS50191; CRAL_TRIO; 1.
 PROSITE; PS00741; DH_1; FALSE_NEG.
 PROSITE; PS50010; DH_2; 2.
 PROSITE; PSS0835; IG_LIKE; 1.
 PROSITE; PSS0003; PH_DOMAIN; 2.
 PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 PROSITE; PS50002; SH3; 1.
 Guanine-nucleotide releasing factor; Phosphorylation; Repeat;
 Transferase; Serine/threonine-protein kinase; ATP-binding;
 Immunoglobulin domain; SH3 domain; Alternative splicing.
 DOMAIN 6 151
 CRAL-TRIO.

FT	REPEAT	252	359	SPECTRIN 1.
FT	REPEAT	479	585	SPECTRIN 2.
FT	REPEAT	819	925	SPECTRIN 3.
FT	REPEAT	1050	1157	SPECTRIN 4.
FT	DOMAIN	1233	1408	DH 1.
FT	DOMAIN	1421	1532	PH 1.
FT	DOMAIN	1597	1653	SH3.
FT	DOMAIN	1910	2086	DH 2.
FT	DOMAIN	2098	2212	PH 2.
FT	DOMAIN	2626	2716	IG-LIKE C2-TYPE.
FT	DOMAIN	2737	2993	PROTEIN KINASE.
FT	DOMAIN	656	659	POLY-GLN.
FT	DOMAIN	1786	1791	POLY-SER.
FT	DOMAIN	2233	2253	POLY-GLY.
FT	DOMAIN	2486	2492	POLY-SER.
FT	DISULFID	2637	2700	POTENTIAL.
FT	NP BIND	2743	2751	ATP (BY SIMILARITY).
FT	BINDING	2766	2786	ATP (BY SIMILARITY).
FT	ACT SITE	2856	2866	BY SIMILARITY.
FT	VARSPLIC	2242	2308	GGGGGGGGGAGGPGSGHGGGPGSGCGAPSTSRSPRSR PGPVRRHPVLVSSAASOAAADKMS -> AAAGVGAAGAA GPVRAATAVAPAAAGAPACAGGPGGSPSLSDITTPCW SLOPPARGQRTRCQ (in isoform 2). /FtId=VSP.004467. Missing (in isoform 2). /FtId=VSP.004468. E->A: 50% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. T->A: 40% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. N->A: NO CHANGE IN NUCLEOTIDE EXCHANGE ACTIVITY. V->A: 90% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. Q->A: 80% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. R->A: 80% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. T->A: 80% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. K->A: LOSS OF NUCLEOTIDE EXCHANGE ACTIVITY. L->A: 40% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. K->A: NO CHANGE IN NUCLEOTIDE EXCHANGE ACTIVITY. E->A: 30% DECREASE IN NUCLEOTIDE EXCHANGE ACTIVITY. SEQUENCE 3038 AA; 341611 MW; 28620F3B513E874B CRC64;
Query Match	6.8%;	Score 592.5;	DB 1;	Length 3038;
Best Local Similarity	34.5%;	Pred. No. 1.9e-13;		
Matches 145;	Conservative	69;	Mismatches 191;	Indels 15; Gaps 6;
QY	53	PPSMQVETDVQAQTGGTAQFAAIEGGPQPQSVTWYKDSVOLVD-----STRLSQQCEGT 107		
Db	2625	PPFVPIPLSEVTCTETCTVLCRCVGRPKASITWKGPENLTNNDGHYSISYDLGEAT 2684		
QY	108	TYSILVRHVASKDAGVYTCCLAQNTGGQGVLCXAELLVLGGNEDPSEKQSHRKLHGFYEV 167		
Db	2685	---LKIIVGVTTEDDGIYTCIAVNDMGSSASSASLRVLGPG--MDGINVMTWKONFDSFYSE 2739		
QY	168	KEETIGRGVFGVKRVQHKGNKILCAKEPIPLRSRTRAQVREYRDILAAALSHPLVTLGLDQ 227		
Db	2740	VAELGRGRFVSVKCKDKQKTKRAVATKFNKKLMKRQDVTHLQGLQSLQHPLLVGLDIT 2799		
QY	228	FEETKTLILILECSSEELDLRYKGVVTAEEVKVYIQQLVEGLHYLHSHGVHLIDIKP 287		
Db	2800	FETFTSVILVLENDQGRLLDCVWRGSLTSGKIRAHLGVLZGAVRYLHNCRIAHLDLKP 2859		
QY	288	SNILMVHP-AREDIKICDQFAQNTTPAELQFSQSPFVSPPEIQQNPVSEASDINAM 346		
Db	2860	ENILVDESIAKPTIKLADFGDAVOLNTYYIHOGLGNPEFAAPEIILGNPVLSTWSV 2919		

QY 347 GVISYLSLTCSSPPAGESDRATILNVLGRVSNWSPMAHLSEDAKDFIKATLQAPQAR 406
Db 2920 GVLITVLSGVSFPDLDVEETCTNICRLDFSPDDYFKGVSKAKFEFVCLLQEDPAKR 2979
QY 407 PSAAQCLSHPWFLKSPFAEEAHFINTKQLKFLARSRWQ---RSLMSYKSLVWRISPL 463
Db 2980 PSAAALQEQW-LQAGNGRSGVLDTSRLTSFIERRHQNDVPIRSIKNQLSRLPRV 3038
RESULT 2
KMLS_BOVIN
ID KMLS_BOVIN STANDARD; PRT; 1176 AA.
AC Q28824;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myosin light chain kinase, smooth muscle (EC 2.7.1.117) (MLCK)
DE [Contains: Telokin].
GN MILK.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=92203148; PubMed=1284247;
RA Kobayashi H., Inoue A., Mikawa T., Kuwayama H., Hotta Y., Masaki T.,
RA Ebashi S.;
RT "Isolation of cDNA for bovine stomach 155 kDa protein exhibiting
RT myosin light chain kinase activity.";
RL J. Biochem. 112:786-791(1992).
CC -!- FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR
CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE
CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT
CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS.
CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
CC light-chain] phosphate.
CC -!- SUBUNIT: ISOFORM TELOKIN BINDS CALMODULIN.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=At least 3 isoforms, Non-muscle, Smooth-muscle (shown
CC here) and Telokin, are produced by alternative initiation.
CC Isoform Non-muscle is the longest and telokin is a C-terminal
CC section without catalytic activity;
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S57131; AB25794.1; -;
DR PIR; JN0583; JN0583.
DR HSSP; P56276; 1TLK.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig-c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 2.
DR Pfam; PF00069; pkinase; 1.

PRINTS; PR00014; FNTYPEIII.
PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IG2; 2.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SI; 1.
KW transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;
KW Alternative initiation.
FT CHAIN 1 1176
FT MYOSIN LIGHT CHAIN KINASE, ISOFORM
FT SMOOTH-MUSCLE.
FT MYOSIN LIGHT CHAIN KINASE, ISOFORM
FT TELOKIN.
FT FOR ISOFORM TELOKIN.
FT 16 X 12 AA TANDEM REPEATS.
FT 1.
FT 2.
FT 3 (INCOMPLETE).
FT 4.
FT 5.
FT 6.
FT 7.
FT 8.
FT 9.
FT 10.
FT 11.
FT 12.
FT 13.
FT 14.
FT 15.
FT 16.
FT IG-LIKE C2-TYPE 1.
FT IG-LIKE C2-TYPE 2.
FT FIBRONECTIN TYPE-III.
FT PROTEIN KINASE.
FT CALMODULIN-BINDING.
FT IG-LIKE C2-TYPE 3.
FT ATP (BY SIMILARITY).
FT ATP (BY SIMILARITY).
FT BY SIMILARITY.
FT POLY-GLU.
FT SEQUENCE 1176 AA; 128824 MW; F53DC6D4D42D4B97 CRC64;
SQ
Query Match 6.3%; Score 550.5; DB 1; Length 1176;
Best Local Similarity 25.3%; Pred. No. 2.1e-12;
Matches 145; Conservative 91; Mismatches 191; Indels 147; Gaps 10;
Qy 45 SSLPALPG-----PPSMQVTIEDVQQTGTAQFAIISGDPQ 83
Db 468 SSLPPVLGTSDATVKKKPAKTPPKAAMPQIIQFPDQKVRAGESVELFKVAGTQPI 527
Qy 84 SVTWYKDSQVLDSTRSLSQOQEGTYSVLVRHVASKDAGVYTCLAQN----- 130
Db 528 TCTWMKFRKQIQDSEHIKVENSEQSKLTIRAAROEHCQGYTLLENKLSRQAQVNLTV 587
Qy 131 -----TGQVLI-----CK-- 138
Db 588 VKRPPAGTFCASDIRSSSLTSLWSYSGSDGSAVSVEIWDSDKTKWELATCRST 647
Qy 139 -----AELLVLG--GDNEP-----DSEKQ-- 155
Db 648 SFNVQDLLPDREYKFRVRAINVVGTSPEQSELTALGEKPEEPKDEVEVSDDEKPE 707
Qy 156 -----SHRKLHGFYEVKKEIGRGVGFVKRVQHKNGKILCAKFIPLRSRTRQAY 207
Db 708 VDRVTVTNTEGKVSDFDIERLGLSGKFGQVFLRVLEKTKTKIWKAGFKKAYSKEKENI 767
Qy 208 R-ERDILAALSHPLVTGLDQFETRTKLILILELCSSELDRLYRKGV-VTEAEVKVYI 265
Db 768 RQISIMNCLHHPKLVQCVDAFEERANIWMVLEIVSGGELPERIIDDELFELTERSCIKYM 827

QY	266	QQLVEGIIHYLHSHGVLDIKPSNILMWHVPAEZDIKIDFGPAQNITPAELQFQYGSPE	325
Db	828	QKISEGVEYIHKQGVHLDLKPENIMCVNKTGTRIKLIDFGLARLENAAGSLKVLFGTPE	887
QY	326	FVSPETIIQQNPVSPASDIWANGVVISYLSLTCSPFAGESDRATLNLVLEGRVSWSSPMAA	385
Db	888	FVAPEVINYEPIGYATDMWISIGVICYLVSLGSPFPGDNDNETLANVTSATWDFDEDAFD	947
QY	386	HLSDADDFIKATLQAPQAPRPAAGLASHPFLKSPABEAHFHNTKQLKFLARSRWQ	445
Db	948	EISDADDFINLLKDMKWRNLCTQLOHPMLMDTKWMEAKQLSKDKRMKKYARRKWQ	1007
QY	446	RSLMSYKSIILVMSRIPEL-----LRGPPDSP	471
Db	1008	KTGNVAVRAIGLSSMAMISGLGRKSTSGTSP	1041

RESULT 3

```

RABBIT
ID KMLS STANDARD; PRT; 1147 AA.
P29294, Q28729;
01-DEC-1992 (Rel. 24, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Myosin light chain kinase, smooth muscle (EC 2.7.1.117) (MLCK)
[Contains: Telokin].
MYLK.
Oryctolagus cuniculus (Rabbit).
OOS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
MAMMALIA Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_taxid=9986;
[1]
SEQUENCE FROM N.A.
TISUS=Smooth muscle;
MEDLINE=92084694; PubMed=1748666;
Gallagher P.J., Herring B.P., Griffin S.A., Stull J.T.;
"molecular characterization of a mammalian smooth muscle myosin light
chain kinase.";
J Biol. Chem. 266:23936-23944 (1991).
[2]
SEQUENCE OF 993-1147 FROM N.A. (TELOKIN).
MEDLINE=92084695; PubMed=1748667;
Gallagher P.J., Herring B.P.;
"The carboxyl terminus of the smooth muscle myosin light chain kinase
is expressed as an independent protein, telokin.";
J. Biol. Chem. 266:23945-23952 (1991).
-!- FUNCTION: Phosphorylates a specific serine in the N-terminus of a
myosin light chain.
-!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
light-chain] phosphate.
-!- ALTERNATIVE PRODUCTS:
Event-Alternative initiation;
Comment=At least 3 isoforms, Non-muscle, Smooth-muscle (shown
here) and Telokin, are produced by alternative initiation.
Telokin is a C-terminal section with no catalytic activity;
-!- DOMAIN: TELOKIN BINDS CALMODULIN.
-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
-!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
-!- SIMILARITY: Contains 1 fibronectin type III domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
-----
EMBL; M76233; AAA73093.1; -
EMBL; M76234; AAA31408.1; -
EMBL; M76181; AAA31409.1; -
PIR; A41675; A41675.
-----

```

Db 583 WYSSYDGGSAVOSYSVEIWDSDYKMWTELATCRSTFNVDRLLDPREKFRVRAINYG 642
QY 147 DNEP-----DSEKQ-----SHRRKLHSFYEVKEEI 171
Db 643 TSFSPQSELTUWGEKPEKDEVEVSDDEKEPEVDYRTVTNTEQKVSDFYDIERL 702
QY 172 GRGVGFVKVQVHKGKNCILCAAKFIPLRSRTRAQAVRERD-----ILALSHPLVTGL 224
Db 703 GSKGFGVFLVEKTKTWAGKFF-----KAYSAREKENIPAEIGIMNCLHHPKLVQC 756
QY 225 LDQFETKTLILLILECSSEELDLRYKGV-VTEAEVQVYIQQLVEGLYHLSHGVLHL 283
Db 757 VDAFEKANIMVLEIVSGSEELFERIDDFELTERECIKYMRQISEGVYHKGIVHL 816
QY 284 DIRPSNLMVHPAREDKICDFGAQNTPEALQFSQYSPFVSPETIIQQNPVSEASDI 343
Db 817 DLKPEIMCVKNGTKRIKIDFGLARLENAGSLKVLFGTPEFVAVEVINEPISYATDM 876
QY 344 WAGVVISYLSLTSSPAGESBATLLNVLEGRVSWSSPMAHLSEDAKDKATLQAP 403
Db 877 WSGVICYILVSLGSPFNGDNDETANVTSATWDFDEAFDEISDDAKDFISNLLKDM 936
QY 404 QARPSAAQCLSHFWFLKSPAEAHFINTKQLKFLARSQWRSYKSLVMSRISPEL 463
Db 937 KRLDCTQCLQHPMLMKDKTKWEAKLSDKRNKYMARRKWKGTGNVAVRAIGRLSSAMI 996
QY 464 LRGPDPSPSLGVARHLCRTGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 523
Db 997 -----SGLSGKSGSTGS-----PTSPLT----- 1014
QY 524 RPSASLPEAEASERSTEAPA---PPASP 549
Db 1015 --AERLETEEDVQAFLEAVEAEKPHVKP 1041

RESULT 4

1001 HUMAN STANDARD; PRT; 1914 AA.
AC Q15746; Q95796; Q95797; Q95798; Q95799; Q14844; Q16794; Q9UBG5;
AC Q9UIT9;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myosin light chain kinase, smooth muscle and non-muscle isoforms
DE (SC 2.7.1.117) (MLCK) [Contains: telokin (Kinase related protein)
DE (KRP)].
GN MYLK OR MLCK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=umbilical vein endothelial cells;
RX MEDLINE=97304466; PubMed=9160829;
RA Garcia J.G.N., Lazar V.L., Gilbert-McClain L.I., Gallagher P.J.,
RA Verin A.D.;
RT "Myosin light chain kinase in endothelium: molecular cloning and
RT regulation.";
RL Am. J. Respir. Cell Mol. Biol. 16:489-494 (1997).
RN [2]
RN REVISIONS.
RA Brukov K.G., Garcia J.G.N.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3A; 3B AND 4).
RC TISSUE=umbilical vein;
RX MEDLINE=99216419; PubMed=10198165;
RA Lazar V.L., Garcia J.G.N.;
RT "A single human myosin light chain kinase gene (MLCK; MYLK).";
RL Genomics 57:256-267 (1999).
RN [4]

RP REVISIONS (ISOFORM 2).
RA Brukov K.G., Garcia J.G.N.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 923-1914 FROM N.A.
RC TISSUE=hippocampus;
RX MEDLINE=96121365; PubMed=8575746;
RA Potier M.-C., Chelot E., Pekarsky Y., Gardiner K., Rossier J.,
RA Turnell W.G.;
RT "The human myosin light chain kinase (MLCK) from hippocampus:
RT cloning, sequencing, expression, and localization to 3qcen-q21.";
RL Genomics 29:562-570 (1995).
RN [6]
RP SEQUENCE OF 1614-1914 FROM N.A.
RC TISSUE=Lung, and Placenta;
RX MEDLINE=20007838; PubMed=10536370;
RA Watterson D.M., Schavocky J.P., Guo L., Weiss C., Chlenski A.,
RA Shinsky V.P., Van Eldik L.J., Haiech J.;
RT "Analysis of the kinase-related protein gene found at human chromosome
RT 3q21 in a multi-gene cluster: organization, expression, alternative
RT splicing and polymorphic marker.";
RL J. Cell. Biochem. 75:481-491 (1999).
RN [7]
RP SEQUENCE OF 1456-1914 FROM N.A.
RC TISSUE=Placenta;
RX Watterson M.D.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR
CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE
CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT
CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS. CENTRAL
CC DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE
CC EDENIA FORMATION. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO
CC CONTROL THE GROWTH INITIATION OF ASTROCYTIC PROCESSES IN CULTURE
CC AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED
CC BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT
CC IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.
CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
CC light-chain] phosphate.
CC -!- SUBUNIT: ISOFORM TELOKIN BINDS CALMODULIN.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing. Named isoforms=6;
CC Comment-Additional isoforms seem to exist;
CC Name=1; Synonyms=Non-muscle isozyme;
CC IsoId=Q15746-1; Sequence=Displayed;
CC Note-The smooth muscle isozyme and telokin are produced by
CC alternative initiation at Met-923 and Met-1761 of isoform 1;
CC Name=2;
CC IsoId=Q15746-2; Sequence=VSP_004791;
CC Name=3A; IsoId=Q15746-3; Sequence=VSP_004792, VSP_004794;
CC Name=3B;
CC IsoId=Q15746-4; Sequence=VSP_004791, VSP_004792, VSP_004794;
CC Name=4;
CC IsoId=Q15746-5; Sequence=VSP_004792, VSP_004793;
CC Name=Del-1790;
CC IsoId=Q15746-6; Sequence=VSP_004795;
CC Event-Alternative initiation;
CC Comment=3 isoforms, 1/Non-muscle isozyme (shown here),
CC smooth-muscle isozyme and telokin, are produced by alternative
CC initiation at Met-1, Met-923 and Met-1761. Telokin has no
CC catalytic activity;
CC -!- TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOZYMES ARE
CC EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN
CC CULTURED ENDOTHELIUM WITH QUALITATIVE EXPRESSION APPEARING TO BE
CC NEITHER TISSUE- NOR DEVELOPMENT-SPECIFIC. NON-MUSCLE ISOFORM 2 IS
CC THE DOMINANT SPICE VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN
CC HAS BEEN FOUND IN A WIDE VARIETY OF ADULT AND FETAL TISSUES.
CC -!- PTM: MLCK IS PROBABLY DOWN-REGULATED BY PHOSPHORYLATION.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -!- SIMILARITY: Contains 9 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U48959; AAC19423.2; -
 DR EMBL; AF069601; AAD15921.2; -
 DR EMBL; AF069602; AAD15922.1; -
 DR EMBL; AF069603; AAD15923.1; -
 DR EMBL; AF069604; AAD15924.1; -
 DR EMBL; X85337; CAA59685.1; -
 DR EMBL; AF096771; AAD51380.1; -
 DR EMBL; AF096766; AAD51380.1; JOINED.
 DR EMBL; AF096767; AAD51380.1; JOINED.
 DR EMBL; AF096768; AAD51380.1; JOINED.
 DR EMBL; AF096769; AAD51380.1; JOINED.
 DR EMBL; AF096770; AAD51380.1; JOINED.
 DR EMBL; AF096774; AAD54018.1; -
 DR EMBL; AF096771; AAD51381.1; -
 DR EMBL; AF096769; AAD51381.1; JOINED.
 DR EMBL; AF096770; AAD51381.1; JOINED.
 DR EMBL; X90870; CAA62378.1; -
 DR HGSP; Q63450; IAO6.
 DR Genew; HGNC.7590; MYLK.
 DR MIM; 600322.
 DR GO; GO:0004687; F-Myosin-light-chain kinase activity; TAS.
 DR GO; GO:0004688; P-protein amino acid phosphorylation; TAS.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00047; ig; 8.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR PROSITE; PSS00835; IG-LIKE; 9.
 DR PROSITE; PSS0107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PSS0108; PROTEIN KINASE ST; 1.
 KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
 KW ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;
 KW Alternative initiation; Alternative splicing.
 FT CHAIN 1 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM NON-MUSCLE ISOZYME.
 FT CHAIN 923 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM
 FT CHAIN 1761 1914 SMOOTH-MUSCLE ISOZYME.
 FT CHAIN 1761 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM
 FT CHAIN 1761 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM
 FT CHAIN 1761 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM
 FT INIT_MET 923 923 FOR ISOFORM SMOOTH-MUSCLE ISOZYME.
 FT INIT_MET 1761 1761 FOR ISOFORM TELOXIN.
 FT DOMAIN 33 122 IG-LIKE C2-TYPE 1.
 FT DOMAIN 161 249 IG-LIKE C2-TYPE 2.
 FT DOMAIN 414 503 IG-LIKE C2-TYPE 3.
 FT DOMAIN 514 599 IG-LIKE C2-TYPE 4.
 FT DOMAIN 620 711 IG-LIKE C2-TYPE 5.
 FT DOMAIN 721 821 IG-LIKE C2-TYPE 6.
 FT DOMAIN 1098 1186 IG-LIKE C2-TYPE 7.
 FT DOMAIN 1238 1326 IG-LIKE C2-TYPE 8.
 FT DOMAIN 1413 1719 FIBRONECTIN TYPE-III.
 FT DOMAIN 1464 1719 PROTEIN KINASE.
 FT DOMAIN 1711 1774 CALMODULIN-BINDING.
 FT DOMAIN 1809 1898 IG-LIKE C2-TYPE 9.
 FT NP_BIND 1470 1478 ATP (BY SIMILARITY).
 FT BINDING 1493 1493 ATP (BY SIMILARITY).
 FT ACT_SITE 1585 1585 BY SIMILARITY.
 FT DOMAIN 1906 1914 POLY-GLU.
 FT DOMAIN 868 998 5 X 28 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 868 998 1-1.
 FT REPEAT 896 923 1-2.

FT REPEAT 924 951 I-3.
 FT REPEAT 952 979 I-4.
 FT REPEAT 980 998 I-5 (INCOMPLETE).
 FT DOMAIN 999 1063 6 X 12 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 999 1003 II-1 (INCOMPLETE).
 FT REPEAT 1004 1015 II-2.
 FT REPEAT 1016 1027 II-3.
 FT REPEAT 1028 1039 II-4.
 FT REPEAT 1040 1051 II-5.
 FT REPEAT 1052 1063 II-6.
 FT VARSPLIC 437 506 VSGIKPENVAMFLSGTPVRRQEGSIEVVEDAGSHVCLLKA
 RTRDGTGYCTASNAQGVSCSWILOVER -> G (in
 isoform 2 and isoform 3B).
 FT FTIC-VSP_004791.
 FT DEVEVSD -> MKWRQQT (in isoform 3A,
 isoform 3B and isoform 4).
 FT FTIC-VSP_004792.
 FT Missing (in isoform 4).
 FT FTIC-VSP_004793.
 FT
 Query Match 6.2%; Score 541; DB 1; Length 1914;
 Best Local Similarity 24.4%; Pred. No. 7.1e-12;
 Matches 140; Conservative 91; Mismatches 196; Indels 146; Gaps 7;
 QY 45 SSLPALPG-----PPSMQVTIEDVQAGTGGTAQFEAIEGDPQP 83
 DB 1208 SSLPVLGTESDATVKKPAPKTPPKAAMPPOITQFEDQKVRAGESVLEFGKVTGTQPI 1267
 QY 84 SVTWKDSVQLVDSTRLSQQQEGTYSIVLRHVASKDAGVYTCIAQNTGGVLCASLLV 143
 DB 1268 TCTWKKFKQIQESEHMKVENSENGSKLTILAAQEHCGCYTLLVENKLGSRQAVLTV 1327
 QY 144 LG----- 145
 DB 1328 VVKPDPAGTPCASDIRSSSLTSLWYSSYDGGSAVQSYSEIWDSSANKTWKELATCRST 1387
 QY 146 -----GDNEP-----DSEKQ--- 155
 DB 1388 SFNVQDLLPDHEYKPRVRAINVTGTSSEPSQSELTVTGKEPEEPKDEVEDDDDEKEPEV 1447
 QY 156 -----SHRRKLGHSFYVKEEIGRGVGFVKRVQHKGNKILCAKTIPLRSRTAAYR 208
 DB 1448 DRTVTINTEQKVSDFDIERLGSRGEGVFRLEKTRKVKWAGKFKVSAKEKENIR 1507
 QY 209 -ERDIALASHPLVTGLDQFETRTKLILILELCSSELDRLRYKGV-VTEARVKYIQ 266
 DB 1508 QEISIMNCLHHPKLVQCVDAFEERKANIWMVLEIVSGGELFERIIDEFELTERECIKYMR 1567
 QY 267 QLVGGLHLSHGVHLHLDIKPSNIMLVHPAREDIKIDFGPAQNTITPAELQFQSYGSPF 326
 DB 1568 QISEGVEIHKQGIIVHLDLKPENINCVNKTGRIKLIDFGLARLENAGSLKVLFGTPEF 1627
 QY 327 VSPETIQNPVSEADIVAMGVISYLSLTCSPSPAGESDRATLLNLEGRVSWSSPMAAH 386
 DB 1628 VAPEVINYEPIGYATDMMSIGVICYLVSGLSPFMGNDNETLANVTSATWDFDEAFDE 1697
 QY 387 LSEDAKDFIKATLQAPQARP-SAAQCLSHPWFLKSPAEAAHFNFKQLKFLARSQR 446
 DB 1688 ISDDAKDFISNLLKDMKNRLDCTCLOHPMLMDKTKNMEAKKLSKDMKMYARRKWK 1747
 QY 447 SLMSYKSLVWRSIPEL-----LRGPPDSP 471
 DB 1748 TGNVRAIYGRUSSAMISGLSKRSKSTGSPSTP 1780
 RESULT 5
 ID_KMLS_CHICK STANDARD; PRT; 1906 AA.
 AC P11799; P19038;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Myosin light chain kinase, smooth muscle and non-muscle isozyms

DE (EC 2.7.1.117) (MLCK) [Contains: Telokin].

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RP [1]

RP SEQUENCE FROM N.A. (ISOFORM MLCK-210).

RX MEDLINE=96033976; PubMed=7589469;

RA Watterson D.M., Collinge M., Lukas T.J., van Eldik L.J., Birukov K.G.,

RA Stepanova O.V., Shirinsky V.P.;

RT "Multiple gene products are produced from a novel protein kinase

RT transcription region".

RL FEBS Lett. 373:217-220(1995).

RN [2]

RN SEQUENCE FROM N.A. (ISOFORM MLCK-108).

RX MEDLINE=90192792; PubMed=2315320;

RA Olson N.J., Pearson R.B., Needleman D.S., Hurwitz M.J., Kemp B.E.,

RA Means A.R.;

RT "Regulatory and structural motifs of chicken gizzard myosin light

RT chain kinase".

RL Proc. Natl. Acad. Sci. U.S.A. 87:2284-2288(1990).

RN [3]

RN SEQUENCE OF 649-1906 FROM N.A., AND PARTIAL SEQUENCE.

RP TISSUE=Fibroblast;

RX MEDLINE=903161738; PubMed=2202734;

RA Shoemaker M.O., Lau W., Shattuck R.L., Kwiatkowski A.P.,

RA Maristany P.E., Guerra-Santos L., Wilson E., Lukas T.J.,

RA van Eldik L.J., Watterson D.M.;

RT "Use of DNA sequence and mutant analyses and antisense

RT oligodeoxynucleotides to examine the molecular basis of nonmuscle

RT myosin light chain kinase autoinhibition, calmodulin recognition, and

RT activity".

RL J. Cell Biol. 111:1107-1125(1990).

RN [4]

RN SEQUENCE OF 1259-1906 FROM N.A.

RP TISSUE=Gizzard;

RX MEDLINE=87157587; PubMed=3030394;

RA Guerrero V. Jr., Russo M.A., Olson N.J., Putkey J.A., Means A.R.;

RA "Domain organization of chicken gizzard myosin light chain kinase

RA deduced from a cloned cDNA".

RL Biochemistry 25:8372-8381(1986).

RN [5]

RN SEQUENCE OF 1750-1906 FROM N.A. (ISOFORM TELOKIN).

RP TISSUE=Gizzard;

RX MEDLINE=93073972; PubMed=1444462;

RA Yoshikawa S., Ikabe M.;

RT "Molecular cloning of the chicken gizzard telokin gene and cDNA".

RL Arch. Biochem. Biophys. 299:242-247(1992).

RN [6]

RN SEQUENCE OF 1750-1906 FROM N.A. (ISOFORM TELOKIN).

RP MEDLINE=92236611; PubMed=1373815;

RA Collinge M., Maristany P.E., Zimmer W.E., Shattuck R.L., Lukas T.J.,

RA van Eldik L.J., Watterson D.M.;

RT "Structure and expression of a calcium-binding protein gene contained

RT within a calmodulin-regulated protein kinase gene".

RL Mol. Cell. Biol. 12:2359-2371(1992).

CC -!- FUNCTION: Phosphorylates a specific serine in the N-terminus of a

CC myosin light chain, which leads to the formation calmodulin/MLCK

CC signal transduction complex which allow selective transduction

CC of calcium signals.

CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin

CC light-chain] phosphate.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative initiation;

CC Comment=At least 3 isoforms, MLCK-210/Non-muscle,

CC MLCK-108/Smooth-muscle and Telokin, are produced by alternative

CC initiation;

CC -!- TISSUE SPECIFICITY: ISOFORM TELOKIN IS EXPRESSED IN GIZZARD,

CC HEART, LUNG, INTESTINE, AND SKELETAL MUSCLE ALTHOUGH THE LEVELS OF

CC THE EXPRESSION IN THE LATTER WERE MUCH LESS THAN THAT IN THE

CC GIZZARD.

CC -!- DOMAIN: TELOKIN BINDS CALMODULIN.

CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.

CC -!- SIMILARITY: Contains 9 immunoglobulin-like C2-type domains.

CC -!- SIMILARITY: Contains 1 fibronectin type III domain.

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X52876; CAA37056.1; -

DR EMBL; X52876; CAA37057.1; -

DR EMBL; X52876; CAA37058.1; -

DR EMBL; M31048; AAA49069.1; -

DR EMBL; M14953; AAA69964.1; -

DR EMBL; M96655; AAA49083.1; -

DR EMBL; M82283; AAA48647.1; -

DR EMBL; M82284; AAB53768.1; -

DR PIR; S68235; S68235.

DR PDB; 1CDL; 31-AUG-94.

DR PDB; 1VRK; 27-APR-99.

DR InterPro; IPR008957; FN III-like.

DR InterPro; IPR003961; FN-III.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR008271; Ser_thr_kinase.

DR InterPro; IPR002290; Ser_thr_kinase.

DR Pfam; PF00047; Ig; 9.

DR Pfam; PF00047; Ig; 9.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00060; FN3; 1.

DR SMART; SM00408; IG2; 8.

DR SMART; SM00220; S_TKc; 1.

DR PROSITE; PS00835; IG_LIKE; 9.

DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;

KW ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;

KW Alternative initiation; 3D-structure.

FT CHAIN 1 1906 MYOSIN LIGHT CHAIN KINASE, ISOFORM MLCK-

FT 210.

FT CHAIN 935 1906 MYOSIN LIGHT CHAIN KINASE, ISOFORM MLCK-

FT 108.

FT CHAIN 1750 1906 MYOSIN LIGHT CHAIN KINASE, ISOFORM

FT TSLOKIN.

FT INIT_MET 935 935 FOR ISOFORM MLCK-108.

FT INIT_MET 1750 1750 FOR ISOFORM TELOKIN.

FT DOMAIN 128 117 IG-LIKE C2-TYPE 1.

FT DOMAIN 156 244 IG-LIKE C2-TYPE 2.

FT DOMAIN 429 517 IG-LIKE C2-TYPE 3.

FT DOMAIN 521 613 IG-LIKE C2-TYPE 4.

FT DOMAIN 637 725 IG-LIKE C2-TYPE 5.

FT DOMAIN 735 830 IG-LIKE C2-TYPE 6.

FT DOMAIN 1084 1172 IG-LIKE C2-TYPE 7.

FT DOMAIN 1225 1313 IG-LIKE C2-TYPE 8.

FT DOMAIN 1330 1400 FIBRONECTIN TYPE-III.

FT DOMAIN 1453 1708 PROTEIN KINASE.

FT DOMAIN 1794 1885 IG-LIKE C2-TYPE 9.

FT NP_BIND 1459 1467 ATP (BY SIMILARITY).

FT BINDING 1482 1482 ATP (BY SIMILARITY).

FT ACT_SITE 1574 1574 BY SIMILARITY.

FT DOMAIN 1716 1728 CALMODULIN AUTOINHIBITION (AM13) REGION

FT (POTENTIAL)

FT DOMAIN 1730 1749 CALMODULIN RECOGNITION (RS20) REGION

FT (POTENTIAL)

FT DOMAIN 1317 1364 MOTIF IA.

FT DOMAIN 1385 1402 MOTIF IB.

FT DOMAIN 660 1833 4 X REPEATS, MOTIF IIA.

FT REPEAT 660 676 IIA-1.
 FT REPEAT 758 774 IIA-2.
 FT REPEAT 1107 1123 IIA-3.
 FT REPEAT 1817 1833 IIA-4 REPEATS, MOTIF IIB.
 FT DOMAIN 693 1866 IIB-1.
 FT REPEAT 693 708 IIB-2.
 FT REPEAT 791 807 IIB-3.
 FT REPEAT 1140 1156 IIB-4.
 FT REPEAT 1281 1297 IIB-5.
 FT REPEAT 1851 1866 IIB-6.
 FT DOMAIN 970 1226 4 X REPEATS, MOTIF III.
 FT REPEAT 970 987 IIB-1.
 FT REPEAT 999 1016 IIB-2.
 FT REPEAT 1061 1078 IIB-3.
 FT REPEAT 1209 1226 IIB-4.
 FT DOMAIN 1700 1763 CALMODULIN-BINDING.
 FT DOMAIN 1896 1906 POLY-GLU.
 FT MOD RES 1748 1748 PHOSPHORYLATION.
 FT MOD RES 1762 1762 PHOSPHORYLATION.
 FT CONFLICT 1439 1439 R -> Q (IN REF. 4).
 SQ SEQUENCE 1906 AA; 210445 MW; AD7D8A3B69EE3363 CRC64;

Query Match 6.1%; Score 532; DB 1; Length 1906;
 Best Local Similarity 20.5%; Pred. No. 1.4e-11;
 Matches 397; Conservative 233; Mismatches 700; Indels 608; Gaps 83;

QY 48 PALPG-PPSMQVIEDVCAQTGCTAQAFAIISGDPOPSVTWKDSVOLVDSTLSQQOEG 106
 DB 21 PSVPAEAPATLPPRNIRVQLGATARFEGKVRGYPEPQITWYRNGHPL-----PEG 71
 QY 107 TTY-----SLVLRHVASKDAGVYTCIAONTGGVCLCAEALLVLGGD-----NEPDS 152
 DB 72 DHYVVDHSIRGIFSLVKGVEGDSKGTCEAANDGGVQVTVELTVGNSLKKYSLPSS 131
 QY 153 EKQSHRKLHSLFYVEKEIGRGVFGVKRVOH-----KGNKILC-----AAK 194
 DB 132 AKTPGGEL-----SVPPVHRPSTINGESPFPKATKNRVVVREGQTR 174
 QY 195 F-PLRSRTAQ-AYRERDILAAHPLVTGLLDQFETKTLILILECSSEELDR-LY 251
 DB 175 FSKKITGRPOQVWTWTKGDI-----HLQNERFNMF-----KTGIQYLEI-QNVQLADAGIY 226
 QY 252 RKGWVTEA-EVKYVIOQLVEGLHYLHSHG-VLHLDIKPSNLMVHPAREDKICDF----- 305
 DB 227 TCTVNSAGKASVSAELTVGGPDKTTHAQPLCMPEKPTL-----ATKAENSDFKOAT 281
 QY 306 --GPAQNI--TPAELOFSQVGSPEFVSPEIIQONPVSEASDIWANGVIVSYLSLTCSSPFA 361
 DB 282 SNGIAKELKSTSLM-----VETKDRLSAKKETFTYSREAKDGKQ 322
 QY 362 GESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLQAPQAPRPSAAQCLSHFWFLKS 421
 DB 323 GQNCQANVPLQESR-GTKGPQVQLKTSITILQAVKAQPEKAEPTT-----FIRQ 374
 QY 422 MPABEAHFINTKOLKFLARSWRORSMSYKSIILVMSRIPELLRPPDSPSLGVARHLCR 481
 DB 375 --AED-----RKRTVQPLMTT-----QENPSL----- 396
 QY 482 DTGSSSSSSSSONELAPARAKSLPPSPVTHSPLHPRGLRPSASLPBBAEASE----- 537
 DB 397 -TQGVSPRSTENRAGVKSVKEEKREPLGLIPQFESR-----PQLEASGGEI 446
 QY 538 -----RSTAPAPAPASPEGAGPPAAQG-----CVRPHSVIRSLFVHOA--GESPEHGALAPGS 588
 DB 447 KFKSKVSGKPKDPVWFKEGVPIKTGEGIQIYEEDGTHCLMLKKAACLDGSGSYSCAAFPN 506
 QY 589 RHPARR-----RHLLKG-----GYTAGALPLGRPLMEHR 619
 DB 507 RGQTSISWLLTVKRPKVEEVAFCPSVSLAGCIVSBQDFVLCYVGGV-----PVEIT 560
 QY 620 -VLEESAAREEQATLLAKAPSEFTALRLPASG---THLAPGHSHLEHDSPT---PRPS 672
 DB 561 WLLNEQPIQYAHSTFEAGVAKLTVDQALPEDDGIYTCIAENNAGRASCQAQVTVKEKSS 620

QY 673 SEACG-EAORLPS--APSGAIPRDMGHFPOGSKQLPSTGGHGTQAQPERPSPDSW----- 725
 DB 621 KKAEGTAAKANKTFAPIFLKGLTDLKVMGDSQVIMT-----VEVSANPCPEIITWLNHG 674
 QY 726 ---GQAPAPCHPKQGSA-----PQEGCSHPAPAVAPCPGSPFPFGSCK-----EAPL 768
 DB 675 KEIQETEDFHFEKKGNEYSLYIQE-----VFPEDTGKY--TCEAWNELGETQTOATL 724
 QY 769 VPSSP-----FLGQOPAPPAPAKASPLDSKMGPGDISLPGRP-----KGPCCSP 814
 DB 725 TVQEPQDGIQPMFISKPRSVTAAGONVLI-----SCAIAAGDFFPTVHFKQGEITP 777
 QY 815 GSASQASSSQ-----VSSLRVGS-----SQVGTEPGPSLD---AEGWTOEADLS 856
 DB 778 GTGCEILQNEIDFTILRNQVSRHAGQYELQIRNQVG-ECSQVSLMLRESSASRAEMLR 836
 QY 857 DSPTPTLQRPQEQVTMRKESLGR--GGYAGVAGYGTFAFGDAGGMLQGQPMWARIANAV 914
 DB 837 DGRESA-----SSGERRDGG--NYGALTFGRTSG-----FKK 866
 QY 915 SOSSE-----EEQEEARAEFSQSEEQEAREASPLQVSPARVPVPEVGRAPTRSPPEPTWED 970
 DB 867 SSSETRAABEEQEDVRGLKRRVETREHTESLRQEAQ----- 906
 QY 971 IGOVSIQVLRDLSGDAEADTISLDISEVDPAYLNLSDLYDIKYLPEEFMIFRKPVPSAQ 1030
 DB 907 -----LDFRDLGKKVSTKFSFB-----DKETPAEQMDFR--ANLQR 943
 QY 1031 PPPSPMAEE-----LAEFPPTWMPGELGPHAGLEITESSESDVALLA 1076
 DB 944 QVKPTLSSEERKVAHQVQDFRSVLAKKGTPTLPEKVPVPPK-AVTD-----F 993
 QY 1077 EAAVGRKRK-----WSPP-----SRSLFHPFGRHLPLDEPAELGLRBRVKASVEHI 1122
 DB 994 RSVLGAKKPPAENGASATPAPNABAGSEAQNATPENSEAPAKPV-VKKEEKNDKCEHG 1052
 QY 1123 SRILKRPGLSKGPPKPKGLASFLUSGLKSWDRAPTFLRELSDETIVLQGVTLACQ 1182
 DB 1053 CAWVDGGIIGKKAENK-----AASKPTPPSKGTAPSTFEKLODAKADGKGLVLOCR 1106
 QY 1183 VSAQPAQAATWSKCAPLESSRSRLIS--ATLKNFQLLTILVVAEDLVGVTVCSVSNALG 1240
 DB 1107 ISSDPASVSWTLDSKAIKSSKISVISQEGTLC--LTIEKWPEDEGGEKCIENANAG 1163
 QY 1241 -----TVTTTGVLRKAER-----PSSSP 1258
 DB 1164 KAECACKVLVEDTSTKAAPAEKTKKPTTLPVLSTESSEATVKKKPAKTPPKAAT 1223
 QY 1259 CPDICEVYADGVLVWKPVSYSYGPV----- 1283
 DB 1224 PPQITQFPEDRKVRAGESVELFAKVGTAPITCTTWMKFKQIQENEVIKIENASNSKLT 1283
 QY 1284 -----TYIVQ-----CS 1290
 DB 1284 ISSTKQBHCYCYTLVVENKLSGROAQVNLTVYDKPDPAGTPPCASDIRSSSLTSLWSYGS 1343
 QY 1291 LEGGS-----WTLASIDFDCCLYS-----KLSRGGTYYTFTACVSKAG 1330
 DB 1344 YDGGSAVQSYTVEIWNVSVDNKNWTDLT-----CRSTSFNVQDLQADREYKFRVRAANVYG 1398
 QY 1331 MGYPSPSEVLLG--GFSHLASEE-----ESQGRSAQPLPSTKT-----FAFQ 1374
 DB 1399 ISEPSQSEVWVKVQKBEELKEEELSDDEGKETEVNRYTIVTINTEKQVSDVYNTIEER 1458
 QY 1375 IQRGFSVVROQCEKASGALAAKII-PYHPKDKTAVLRVEALKGLRHPHQAALHAAYL 1433
 DB 1459 LSGKFGQVFRIVEKTKVWAGKFFKAYSASAKENIRDEISINWCLHHPKLQCVDAFE 1518
 QY 1434 SPRHLVILIELCSGPPELLP-CLAERASYSSEVKDYLMQMLSATQYLNQHILHLDLRSE 1492
 DB 1519 EKANTVMVLEWVGSGELFERIIDEFELTERECIKYMRQISEGVETIHKQIIVHLDLRKE 1578

1493	QY	NNIITEY--NLLKVDLGHNAQSQB---KULPSDKFYDLTMAPELLSSGQAVPQTDI	1547
1494	QY		
1495	QY		
1496	QY		
1497	QY		
1498	QY		
1499	QY		
1500	QY		
1501	QY		
1502	QY		
1503	QY		
1504	QY		
1505	QY		
1506	QY		
1507	QY		
1508	QY		
1509	QY		
1510	QY		
1511	QY		
1512	QY		
1513	QY		
1514	QY		
1515	QY		
1516	QY		
1517	QY		
1518	QY		
1519	QY		
1520	QY		
1521	QY		
1522	QY		
1523	QY		
1524	QY		
1525	QY		
1526	QY		
1527	QY		
1528	QY		
1529	QY		
1530	QY		
1531	QY		
1532	QY		
1533	QY		
1534	QY		
1535	QY		
1536	QY		
1537	QY		
1538	QY		
1539	QY		
1540	QY		
1541	QY		
1542	QY		
1543	QY		
1544	QY		
1545	QY		
1546	QY		
1547	QY		
1548	QY		
1549	QY		
1550	QY		
1551	QY		
1552	QY		
1553	QY		
1554	QY		
1555	QY		
1556	QY		
1557	QY		
1558	QY		
1559	QY		
1560	QY		
1561	QY		
1562	QY		
1563	QY		
1564	QY		
1565	QY		
1566	QY		
1567	QY		
1568	QY		
1569	QY		
1570	QY		
1571	QY		
1572	QY		
1573	QY		
1574	QY		
1575	QY		
1576	QY		
1577	QY		
1578	QY		
1579	QY		
1580	QY		
1581	QY		
1582	QY		
1583	QY		
1584	QY		
1585	QY		
1586	QY		
1587	QY		
1588	QY		
1589	QY		
1590	QY		
1591	QY		
1592	QY		
1593	QY		
1594	QY		
1595	QY		
1596	QY		
1597	QY		
1598	QY		
1599	QY		

```

RESULT 6
-----
DAX3 MOUSE
ID DAX3 MOUSE STANDARD; PRT; 448 AA.
AC 054784;
15-MAR-2004 (Rel. 43, Created)
15-MAR-2004 (Rel. 43, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
DEATH-associated protein kinase 3 (EC 2.7.1.37) (DAP kinase 3) (DAP-
like kinase) (Dlk) (Zip-Kinase).
DAPK3 OR ZIPK.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, INTERACTION WITH
ATP4, SUBCELLULAR LOCATION, AND MUTAGENESIS OF LYS-42; VAL-422;
VAL-429 AND LEU-436.
MEDLINE=98147805; PubMed=9488481;
Kawai T., Matsumoto M., Takeda K., Sanjo H., Akira S.;
"ZIP-Kinase, a novel serine/threonine kinase which mediates
apoptosis.";
Mol. Cell. Biol. 18:1642-1651(1998).
[2]
SUBCELLULAR LOCATION, AND INTERACTION WITH DAXX AND PAWR.
MEDLINE=22799111; PubMed=12917339;
Kawai T., Akira S., Reed J.C.;
"ZIP kinase triggers apoptosis from nuclear PML oncogenic domains.";
Mol. Cell. Biol. 23:6174-6186(2003).
-!- FUNCTION: Serine/threonine kinase which acts as a positive
regulator of apoptosis.
-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
-!- COFACTOR: Magnesium.
-!- SUBUNIT: Homodimer or forms heterodimers with ATP4. Both
interactions require an intact leucine zipper domain and
oligomerization is required for full enzymatic activity. Also
binds to DAXX and PAWR, possibly in a ternary complex which plays
a role in caspase activation.
-!- SUBCELLULAR LOCATION: Nuclear. Relocates to the cytoplasm on
binding PAWR where the complex appears to interact with actin
filaments.
-!- TISSUE SPECIFICITY: Highly expressed in heart, brain, lung,
skeletal muscle, kidney and testis. Lower levels in liver and
spleen.
-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
kinase subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch.
-----
EMBL; AB007143; BAA24954.1; -.
HSSP; C63450; IA06.
MGD; MGI:1203520; DapK3.
InterPro; IPR000719; Prot kinase.
InterPro; IPR002290; Ser Thr kinase.

```

DR	InterPro; IPR008271; Ser_thr_pkin_AS.
DR	Pfam; PF00069; pkinase; 1.
DR	ProDom; PD00001; Prot kinase; 1.
DR	SMART; SM00220; S_TK; 1.
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR	PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding; Apoptosis;
FT	Nuclear protein.
FT	DOMAIN 13 275 PROTEIN KINASE.
FT	NP_BIND 19 27 ATP (BY SIMILARITY).
FT	BINDING 42 42 ATP.
FT	ACT_SITE 139 139 BY SIMILARITY.
FT	MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
FT	MUTAGEN 422 422 V->A: DECREASED ACTIVITY; WHEN ASSOCIATED
FT	MUTAGEN 429 429 WITH A-429 AND A-436.
FT	MUTAGEN 436 436 V->A: DECREASED ACTIVITY; WHEN ASSOCIATED
FT	MUTAGEN 436 436 L->A: DECREASED ACTIVITY; WHEN ASSOCIATED
FT	SEQUENCE 448 AA; 51421 MW; DA32EF3EB1F20EFC CRC64;
QY	Query Match 5.6%; Score 489; DB 1; Length 448;
DB	Best Local similarity 38.1%; Pred. No. 1,1e-10;
DB	Matches 114; Conservative 56; Mismatches 117; Indels 12; Gaps 5;
QY	165 YEKVEISGRGVGVKRVQHKNTLKCAAKTPIR---SR---TRAQAVREDIIAALS 217
DB	13 YEMGEELSGQPAIVRAKCOQKGTGMEYAAKTKRRLPSSRRGVSRREIYREVSILREIR 72
QY	218 HPVATGLDDQFPRKTLIIILELSCSEBLLRLRVKGVVTAEBVKVYIQQLVEGLHYLHS 277
DB	73 HPNIIITHDVFNKTDVLLIELVSGGELFDFAEKESLTEDATQFLKQILDGVHYLHS 132
QY	278 HGVHLDDIKPSNLMV--HPAREDITKIDCFQAGNITPAELQFSQYGSFVSPETIIQQN 335
DB	133 KRAHFDLPENIMLIDKRAASPRKILDFGAHRIEAGSEFKNIFGTPEFVAPEIYNYE 192
QY	336 PVSEASDIWAMGVISYLSLTCSPFAGESDRATLLANLEGRVSWSSPMAHLSEDADKFI 395
DB	193 PLGLEADWMSIGVITVILLGSGAPFLGTQKOTLTNISAVNYDDEYFSGTSELAKDFI 252
QY	396 KATLQAPQARSAACLSHPFLKSPAEAAHFNITKQLFLIARSRWORSIMSVKSI 454
DB	253 RRLVADPKRRMTIAQLSHVWKVRRDGAQKPRRRL--AARLR-RYSLSKSHSM 308
RESULT 7	
DAK3_RAT	STANDARD; PRT; 448 AA.
ID_DAK3_RAT	
AC_O88764;	
DT_15-MAR-2004	(Rel. 43, Created)
DDT_15-MAR-2004	(Rel. 43, Last sequence update)
DEDT_15-MAR-2004	(Rel. 43, Last annotation update)
DE	Death-associated protein kinase 3 (EC 2.7.1.37) (DAP kinase 3) (DAP-
DE	like kinase) (Dlk) (ZIP-kinase).
GN_DAPK3 OR ZIPK.	
OS_Rattus norvegicus (Rat).	
OC_Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC_Mammalia; Eutharia; Rodentia; Sciurognathi; Muridae; Rattus.	
NCBI_TaxID=10116;	
[1]	
RN	SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND SUBCELLULAR
RRP	LOCATION.
RRP	
RRR_MEDLINE=99054663; PubMed=9840288;	
RRR_Kogel D., Pictner O., Landeberg G., Christian S., Scheidtmann K.H.;	
RRR_TRT "Cloning and characterisation of Dlk, a novel serine/threonine kinase	
RRR_TRT that is tightly associated with chromatin and phosphorylates core	
RRR_TRT histones.";	
RRR_ONcogene 20:2645-2654 (1998).	
[2]	
RRR_FUNCTION, INTERACTION WITH ATF4 AND PAWR, SUBCELLULAR LOCATION, AND	
RRR_MUTAGENESIS OF LYS-42.	

AC	P53355; Q9BTL8;
AD	01-OCT-1996 (Rel. 34, Created)
DT	15-MAR-2004 (Rel. 43, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Death-associated protein kinase 1 (EC 2.7.1.37) (DAP kinase 1).
DE	DAPK1 OR DAPK
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A., FUNCTION, INDUCTION, AND MUTAGENESIS OF LYS-42.
RP	MEDLINE=95129831; PubMed=7828849;
RX	Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT	"Identification of a novel serine/threonine kinase and a novel 15-kD
RT	protein as potential mediators of the gamma interferon-induced cell
RT	death";
RT	Genes Dev. 9:15-30(1995).
RL	[2]
RN	REVISIONS TO 164-171.
RP	Feinstein E.;
RP	Submitted (APR-1997) to the EMBL/GenBank/DBSJ databases.
RL	[3]
RN	SEQUENCE OF 1-367 FROM N.A.
RP	Kanine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RP	Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA	Phelan M., Farmer A.;
RA	"Cloning of human full-length CDSs in BD Creator(TM) system donor
RT	vector";
RT	Submitted (MAY-2003) to the EMBL/GenBank/DBSJ databases.
RL	[4]
RN	FUNCTION, AND SUBCELLULAR LOCATION.
RP	MEDLINE=20094983; PubMed=10629061;
RX	Inbal B., Shani G., Cohen O., Kissil J.I., Kimchi A.;
RT	"Death-associated protein kinase-related protein 1, a novel
RT	serine/threonine kinase involved in apoptosis";

[MOL. CELL. BIOL. 20:1044-1054(2000)].
 FUNCTION, ENZYME REGULATION, MUTAGENESIS OF LYS-42; SER-308 AND
 SER-313, AND PHOSPHORYLATION OF SER-308.
 PubMed=11579085;
 Shohat G., Spiwak-Kroizman T., Cohen O., Bialik S., Shani G.,
 Berissi H., Eisenstein M., Kimchi A.;
 "The pro-apoptotic function of death-associated protein kinase is
 controlled by a unique inhibitory autophosphorylation-based
 mechanism.";
 J. Biol. Chem. 276:47460-47467(2001).
 -|- FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase
 which acts as a positive regulator of apoptosis.
 -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 -|- COFACTOR: Magnesium.
 -|- ENZYME REGULATION: Negatively regulated by autophosphorylation on
 Ser-308.
 -|- SUBCELLULAR LOCATION: Cytoplasmic. Colocalizes with the actin
 filament system.
 -|- INDUCTION: Up-regulated following treatment with interferon-gamma.
 -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
 Kinase subfamily.
 -|- SIMILARITY: Contains 10 ANK repeats.
 -|- SIMILARITY: Ref.1 sequence differs from that shown due to
 framehifts in positions 466 and 477.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; X76104; CAA53712.1; ALT FRAME.
 EMBL; BT006935; AAP35581.1; ALT TERM.
 PDB; 1UKT; 01-APR-02

CC-----
 DDR ENBL; X76104; CAAS3712.1; ALT FRAME.
 DDR ENBL; BT006935; AAP35581.1; ALT TERM.
 DDR PDB; 1JKT; 01-APR-02.

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numa K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,
RA Verardo R., Wagner L., Wymshaw-Boris A., Yanagisawa M., Yang L., Wells C.,
RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang L., Wells C.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume M., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Landier E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
CC -!- FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase.
CC Isoform 1 is a negative regulator of apoptosis in direct contrast
CC to its human homolog. Isoform 2 has no effect on apoptosis and its
CC function is unknown.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium.
CC -!- ENZYME REGULATION: Negatively regulated by autophosphorylation on
CC Ser-308 (by similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonym=Beta;
CC IsoId=Q80Y87-1; Sequence=Displayed;
CC Name=2; Synonym=Alpha;
CC IsoId=Q80Y87-2; Sequence=VSP_050629;
CC -!- TISSUE SPECIFICITY: High levels in bladder, uterus, vas deferens,
CC lung, liver and kidney.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
CC kinase subfamily.
CC -!- SIMILARITY: Contains 10 ANK repeats.
CC -!- SIMILARITY: Contains 1 death domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY245540; AA091934.2; -;
DR EMBL; AY245541; AA091935.1; -;
DR EMBL; X97048; CA65762.1; -;
DR EMBL; BC021490; AAH21490.1; ALT_INIT.
DR EMBL; BC026671; AAH26671.1; -;
DR EMBL; BC057317; AAH57317.1; -;
DR EMBL; BC060161; AAH60161.1; -;
DR EMBL; AK031533; BAB28681.1; -;
DR MGD; MGI:1916885; Dapki.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00023; ank; 8.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01415; ANKYRIN.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK_8.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PSS0088; ANK_REPEAT; 7.
DR PROSITE; PSS0017; DEATH_DOMAIN; 1.
DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PSS0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PSS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW Alternative splicing.
FT DOMAIN 13 275 PROTEIN_KINASE.
FT DOMAIN 267 334 CALMODULIN-BINDING (BY SIMILARITY).
FT REPEAT 378 407 ANK 1.
FT REPEAT 411 440 ANK 2.
FT REPEAT 444 473 ANK 3.
FT REPEAT 477 506 ANK 4.
FT REPEAT 510 539 ANK 5.
FT REPEAT 543 572 ANK 6.
FT REPEAT 576 605 ANK 7.
FT REPEAT 609 638 ANK 8.
FT REPEAT 675 704 ANK 9.
FT REPEAT 1164 1196 ANK 10.
FT DOMAIN 1312 1396 DEATH.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 139 139 BY SIMILARITY.
FT MOD_RES 308 308 PHOSPHORYLATION (AUTO-) (BY
FT SIMILARITY).
FT VARSPIC 1431 1442 Missing (in isoform 2).
FT CONFLICT 354 354 /FTID=VSP_050629.
FT CONFLICT 441 441 D -> N (IN REF. 2).
FT CONFLICT 461 461 K -> Q (IN REF. 3).
FT CONFLICT 960 960 AAH57317/AAH60161.
FT CONFLICT 1000 1000 V -> A (IN REF. 2).
FT CONFLICT 1038 1042 S -> P (IN REF. 2).
FT CONFLICT 1103 1104 N -> T (IN REF. 2).
FT CONFLICT 1105 1105 RWLCT -> PMALH (IN REF. 2).
FT CONFLICT 1105 1105 A -> V (IN REF. 1).
SQ SEQUENCE 1442 AA; 161440 MW; 243A14D7C6598F63 CRC64;
Query Match 5.4%; Score 469; DB 1; Length 1442;
Best Local Similarity 33.5%; Pred. No. 1.6e-09;
Matches 110; Conservative 64; Mismatches 128; Indels 26; Gaps 5;
Qy 158 RKLLHSEYVEKEIGEGVGFVKRVQHKNTLCRAKETPLR----SR---TRAQAYRER 210
Db 6 QENVDDYDTGELGSGQFAVVKCKRKSTGLQYAAKFKKRTKSSRRGVSREDIDREV 65
Qy 211 DILAALSHPLVTGLDQFETKRLTILILELCSSEELDLRYKGVVTEAEVYVYQQLVE 270
Db 66 SILKEIRHENVITLHEVYENKTDVILILELVAGGELDFDLAEKESLTEEEATEFLKQILS 125
Qy 271 GLYLSHSGVHLIDIKPSNLMV--HPAREDIKLCDFGAQNIIPAEQPSQSGSPPEFVS 328
Db 126 GVYLSLQIAHFDLKPENIMLLDRNVKPKIKIIDFGLAHKIDFGNEFKNIFGTPEFVA 185
Qy 329 PEIQONPVSEASDIWAMGVISYLSLTCSPPFAGESDRATLLNVLEGRVSWSSPMAAHL 388
Db 186 PEIVNYEPLGLEADNWSIGVITYILLGASPPFLGDTKQETLANVSAVNYDPEEEFFNTS 245
Qy 389 EDAKDPKATLQAPQAPSAQCSLHPWF---LKSMPAEAEHFINTKQLKFLALASRW 444
Db 246 TLAKQFIRLLVDPKRWTKMTQDSLQHPWIKPKTKQALSRKASAVNWEKPKFAARKKW 305
Qy 445 Q-----RSLMSYKSIILMRS 459
Db 306 KQSVRLISLCQLRSRSLSRNMSVARS 333
RESULT 11
ID_KML2_HUMAN STANDARD; PRT; 595 AA.
AC Q9H1R3; Q96184;
DT 28-FEB-2003 (Rel. 41, Created)

[illegible]

QY 43 LHSSLPALPGPPSMQVTTIEDVQQTGGTAQFAIEG-----DPPPSVTWYKDSVQL 94

Db 146 LHS-----PSCPAIISSEKILAKPPSEASELTFEGVPMTHSTDPRAKAEKNI-L 199
Qy 95 VDSTR-----LSQQEGTYSILVLRHASKDAGVYVTCIAQNTGGQVLC-----KA 139
Db 200 AESQKEVEKTPGQAGAKMGDTSGICLFOAVSEKSEV-----GQALCLTAREE 250
Qy 140 ELLVLGGNEDSEKSHRR-----KLHSFYEV--KEETGRGVFGFKVRVQHGKNGKILC 191
Db 251 DCFOTLDCPPPPAPPPFPRMVELRTGNVSSEFSMNSKEALGGGKFGAVCTCMKATGLKL 310
Qy 192 AAKRTPLRS--TRAQAYRERDILAAALSHPLVTGLLDQDPETRTKLILILELCSSEELDLRL 250
Db 311 AAKVJKGTPOKEMVLEIEVMQNLNRNLIQLYAAIETHEIVLFWIEGGELEPERI 370
Qy 251 YRKGV-VTEAEVKYVIOQLVGLHSHGVHLHDIKPSNLMLVHPAREDIKICDFGFAQ 309
Db 371 VDEDVHLTEVDTWVFRQICDGLFMKQVRVHLDLKPENLICVNTTGHVVKIIDFGILAR 430
Qy 310 NITPAELQFSQGSSEFVSPEIQQNPVSEASDIWAMGVISYLSITCSPSPAGESDRATL 369
Db 431 RYNPEKLVNFGTPEFVSPEVNVYDQISDKTDMWMSGVITIMLSGLSPFLGDDDTETL 490
Qy 370 LNVLEGRVSWSPMAHLSEDAKPIKATLQAPQAPRSPAAQCLSHWFLKSPAEBAHF 429
Db 491 NNVLSGNVYFDEETEAYSDEAKFVSNLIVKQDQARNNAQCLAHWP-LNNL-AEKAKR 548
Qy 430 INTK-----QLKFLARSRWORSLS 450
Db 549 CNRLKSOQLKKYLMKRWKKNFIA 574

RESULT 12

DAK2 HUMAN
ID DAK2 HUMAN STANDARD; PRT; 454 AA.
AC O43293;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Death-associated protein kinase 3 (EC 2.7.1.37) (DAP kinase 3) (DAP-
like kinase) (DLK) (ZIP-Kinase).
GN DAPK3 OR ZIPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98147805; PubMed=9488481;
RA Kawai T., Matsumoto M., Takeda K., Sanjo H., Akira S.;
RT "ZIP kinase, a novel serine/threonine kinase which mediates
apoptosis".
RL Mol. Cell. Biol. 18:1642-1651(1998).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=99283879; PubMed=10356987;
RA Murata-Hori M., Suizu F., Iwasaki T., Kikuchi A., Hosoya H.;
RT "ZIP kinase identified as a novel myosin regulatory light chain kinase
in HeLa cells".
RL FEBS Lett. 451:81-84(1999).
RN [3]
RP FUNCTION, AND INTERACTION WITH DAXX AND PAWR.
RX MEDLINE=22799111; PubMed=12917339;
RA Kawai T., Akira S., Reed J.C.;
RT "ZIP kinase triggers apoptosis from nuclear PML oncogenic domains".;
RL Mol. Cell. Biol. 23:6174-6186(2003).
CC -I- FUNCTION: Serine/threonine kinase which acts as a positive
regulator of apoptosis.
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- COFACTOR: Magnesium.
CC -I- SUBUNIT: Homodimer or forms heterodimers with ATF4. Both
interactions require an intact leucine zipper domain and
oligomerization is required for full enzymatic activity. Also

CC binds to DAXX and PAWR, possibly in a ternary complex which plays
a role in caspase activation.
CC -I- SUBCELLULAR LOCATION: Nuclear. Relocates to the cytoplasm on
binding PAWR where the complex appears to interact with actin
filaments (By similarity).
CC -I- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
kinase subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC ENBL; AB007144; BAA24955.1; -;
DR ENBL; AB022341; BAA81746.1; -;
DR HSSP; Q63450; 1A06.
DR Genew; HGNC:2676; DAPK3.
DR MIM; 603289; -;
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Apoptosis;
KW Nuclear protein.
KW DOMAIN 13 275 PROTEIN KINASE
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 139 139 BY SIMILARITY.
SQ SEQUENCE 454 AA; 52535 MW; 5677308A6A1CF0 CRC64;

Query Match 5.3%; Score 466.5; DB 1; Length 454;
Best Local Similarity 38.9%; Pred. No. 6.4e-10;
Matches 102; Conservative 52; Mismatches 99; Indels 9; Gaps 3;
Qy 165 YEVKEEIGRGVFGFKVRVQHGKNGKILCAKFIPLR-----SR---TRAQAYRERDILAAALS 217
Db 13 YEMGEELGSGQFAIVRCRQKGTGKQKVEYAKFIKKRLLSSRRGVSREIEREVNILEIR 72
Qy 218 HPLVTGLLDQFETRTKLILILELCSSEELDLRLVKGWVTEAEVKYVIOQLVGLHSHLS 277
Db 73 HPNITLHDFENKTDVVLILELVSGGELDFLAEKESLDEATQFLKILDGVHLS 132
Qy 278 HGVHLHDIKPSNLMLV--HPAREDIKICDFGFAQNTIPPAELQFSQGSSEFVSPEIQQN 335
Db 133 KRIAHFDLKPENIMLLQKNVNPRIKLIDFGIAHKIEAGNEFKNIFGTPEFVAPEIYNYE 192
Qy 336 PVSEASDIWAMGVISYLSITCSPSPAGESDRATLNVLEGRVSWSPMAHLSEDAKDFI 395
Db 193 PLGLEADMWSIGVITYILLGASPLGLTQETLTNISAVNYDFEYFNTSLAKDFI 252
Qy 396 KATLQAPQAPRSPAAQCLSHWP 417
Db 253 RRLIVKDPKRWMTIAQSLEHWS 274

RESULT 13

DAK2 HUMAN
ID DAK2 HUMAN STANDARD; PRT; 370 AA.
AC Q9UIK4; O75892;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

15-MAR-2004 (Rel. 43, Last annotation update)
 DE Death-associated protein kinase 2 (EC 2.7.1.37) (DAP kinase 2) (DAP-kinase)
 GN DAPK2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1].
 RN SEQUENCE FROM N.A., FUNCTION, ENZYME REGULATION, TISSUE SPECIFICITY,
 RP AND MUTAGENESIS OF LYS-52.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=9303019; PubMed=10376525;
 RA Kawai T., Nomura F., Hoshino K., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Akira S.;
 ET "Death-associated protein kinase 2 is a new calcium/calmodulin-
 RT dependent protein kinase that signals apoptosis through its catalytic
 RT activity."; Mol. Cell. Biol. 20:1044-1054(2000).
 RL Oncogene 18:3471-3480(1999).
 [2].
 RN SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, HOMODIMERIZATION,
 RP AND MUTAGENESIS OF LYS-52.
 RC TISSUE=Kidney;
 RX MEDLINE=20094983; PubMed=10629061;
 RA Inbal B., Shani G., Cohen O., Kissil J.L., Kimchi A.;
 ET "Death-associated protein kinase-related protein 1, a novel
 RT Serine/threonine kinase involved in apoptosis."; Mol. Cell. Biol. 20:1044-1054(2000).
 RL EMO J. 20:1099-1113(2001).
 [3].
 RN FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase
 CC which acts as a positive regulator of apoptosis.
 CC -| CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -| COFACTOR: Magnesium.
 CC -| ENZYME REGULATION: Negatively regulated by autophosphorylation on
 CC Ser-318.
 CC -| SUBUNIT: Homodimer. Homodimerization is required for apoptotic
 CC function and is inhibited by autophosphorylation at Ser-
 CC 318.
 CC -| SUBCELLULAR LOCATION: Cytoplasmic.
 CC -| TISSUE SPECIFICITY: Ubiquitously expressed in all tissue types
 CC examined. High levels in heart, lung and skeletal muscle.
 CC -| SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
 CC kinase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB018001; BRA8063.1; -;
 CC EMBL; AF052941; AAC3500.1; -;
 CC HSP; O63450; IAO6.
 CC Genew; HGNC:2675; DAPK2.
 CC InterPro; IPR000719; Prot kinase.
 CC InterPro; IPR002290; Ser_Thr_pkinase.
 CC InterPro; IPR008271; Ser_Thr_pkin_AS.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot kinase; 1.
 CC SMART; SMC0220; S_TKc; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
 Phosphorylation; ATP-binding; Apoptosis.
 KW DOMAIN 23 285 PROTEIN KINASE.
 FT 277 344 CALMODULIN-BINDING.
 FT NP_BIND 29 37 ATP (BY SIMILARITY).
 FT BINDING 52 52 BY SIMILARITY.
 FT ACT_SITE 149 149 K->A: LOSS OF ACTIVITY, APOPTOTIC
 FT MOD_RES 318 318 FUNCTION, AND OF AUTOPHOSPHORYLATION.
 FT MUTAGEN 52 52 S->A: NO EFFECT ON CA(2+)-CALMODULIN
 FT 299 299 INDEPENDENT PHOSPHORYLATION OR APOPTOTIC
 FT 299 299 ACTIVITY.
 FT 299 299 BINDING: LOSS OF CA(2+)-CALMODULIN
 FT 299 299 BINDING; INCREASE IN ACTIVITY, LOSS OF
 FT 299 299 AUTOPHOSPHORYLATION.
 FT 318 318 S->A: LOSS OF CA(2+)-CALMODULIN
 FT 318 318 INDEPENDENT PHOSPHORYLATION, INCREASE IN
 FT 318 318 APOPTOTIC ACTIVITY.
 FT 318 318 S->D: ABOLISHES APOPTOTIC ACTIVITY.
 FT 320 320 S->A: NO EFFECT ON CA(2+)-CALMODULIN
 FT 320 320 INDEPENDENT PHOSPHORYLATION OR APOPTOTIC
 FT 323 323 ACTIVITY.
 FT 323 323 S->A: NO EFFECT ON CA(2+)-CALMODULIN
 FT 323 323 INDEPENDENT PHOSPHORYLATION OR APOPTOTIC
 FT 329 329 ACTIVITY.
 FT 329 329 T->A: NO EFFECT ON CA(2+)-CALMODULIN
 FT 329 329 INDEPENDENT PHOSPHORYLATION OR APOPTOTIC
 FT 241 241 ACTIVITY.
 FT 253 253 A -> S (IN REF. 2).
 FT 253 253 Q -> H (IN REF. 2).
 SQ SEQUENCE 370 AA; 42898 MW; 035E914BBCD861A2 CRC64;
 Query Match 5.3%; Score 465.5; DB 1; Length 370;
 Best Local Similarity 32.4%; Pred. NO. 5.7e-10;
 Matches 121; Conservative 72; Mismatches 133; Indels 47; Gaps 12;
 QY 158 RRLHSHFVKEEIGRGVGVGVKVKHKGKILCAAKFIPLR---SR---TRAQAVRER 210
 DB 16 QKQVEDFDIGSELGGQFAIVKKCKREKSTGLEVAAPFKKQSRASRGVSREREIEV 75
 QY 211 DILAALSHPLVGLDQFETKRLTILILELSSBELDLRLYKGVVTEAEVKKVIOQLVE 270
 DB 76 SILROVLHNVITLHDVYENRTDVVILELVSGGELFELAQKESLEEATSFIKQILD 135
 QY 271 GLHYLHSHGVHLDDIKPSNLMV-----HPAREDIKICDFGAQNTIPALQFSQ-YG 322
 DB 136 GVNLYHTKKIAHFDLKPENIMLLDKNIPH-----IKLIDFGLAHEIDG-VFVKMIFG 189
 QY 323 SPEFVSPEIIQONPVSEASDIWAMGVISYLSLTCSPPAGESDRATLLNVLEGRVSWSSP 382
 DB 190 TFEFVAPEIVNVEPLGLEADMMWSIGVITYILLGASPLFGDTQETLANITAVSYDFDEE 249
 QY 383 MAHLSEDAKDFIKATLQAPQAPSAACCLSHFWEL-----KSPAAEEAHFINTKOLKF 437
 DB 250 FFSQTSSELAQDFIRKLLVYKTKLTQELRHPAIPVDNQQAQVRES-VVNLNFRK 308
 QY 438 LLARSRWORSLSMYKSIIL--VMSIPELLRGPPDPSPLGVARHLCRTGGSSSSSSSDN 495
 DB 309 QTVRRRWKLS-FSIVSLCNHLTRSLMKKVLHURPD-----EDLENCSDET 351
 QY 496 ELAPPARAKSLPP 508
 DB 352 E-EDIARRKALHP 363
 RESULT 14
 XML2_RABIT
 ID XML2_RABIT STANDARD; PRT; 607 AA.
 AC P07313;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Myosin light chain kinase 2, skeletal/cardiac muscle (EC 2.7.1.117)
 DE (MLCK2).
 GN MYLK2.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90110242; PubMed=1688558;
 RA Herring B.P., Stull J.T., Gallagher P.J.;
 RT "Domain characterization of rabbit skeletal muscle myosin light chain
 kinase";
 RL J. Biol. Chem. 265:1724-1730(1990).
 RN [2]
 RP SEQUENCE OF 1-603.
 RX MEDLINE=87101105; PubMed=3542042;
 RA Takio K., Blumenthal D.K., Walsh K.A., Titani K., Krebs E.G.;
 RT "Amino acid sequence of rabbit skeletal muscle myosin light chain
 kinase";
 RL Biochemistry 25:8049-8057(1986).
 RN [3]
 RP SEQUENCE OF 295-603.
 RX MEDLINE=86104095; PubMed=3841288;
 RA Takio K., Blumenthal D.K., Edelman A.M., Walsh K.A., Krebs E.G.,
 RA Titani K.;
 RT "Amino acid sequence of an active fragment of rabbit skeletal muscle
 myosin light chain kinase";
 RL Biochemistry 24:6028-6037(1985).
 RN [4]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=21590235; PubMed=11733062;
 RA Davis J.S., Hassanadeh S., Wintsky S., Lin H., Satorius C.,
 RA Vemuri R., Aletas A.H., Wen H., Epstein N.D.;
 RT "The overall pattern of cardiac contraction depends on a spatial
 gradient of myosin regulatory light chain phosphorylation";
 RL Cell 107:631-641(2001).
 RN [5]
 RP STRUCTURE BY NMR OF 577-602.
 RX MEDLINE=92263094; PubMed=1585175;
 RA Ikura M., Clore G.M., Gronenborn A.M., Zhu G., Klee C.B., Bax A.;
 RT "Solution structure of a calmodulin-target peptide complex by
 multidimensional NMR";
 RL Science 256:632-638(1992).
 CC -!- FUNCTION: Implicated in the level of global muscle contraction and
 cardiac function (by similarity). Phosphorylates a specific serine
 in the N-terminus of a myosin light chain.
 CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
 light-chain] phosphate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Colocalizes with phosphorylated
 myosin light chain (RLCP) at filaments of the myofibrils.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isp-sib.ch).
 CC
 CC EMBL; J05194; AAA31400.1; -;
 CC PIR; A35021; A35021.
 CC PDB; 2BBM; 31-JAN-94.
 CC PDB; 2BBN; 31-JAN-94.
 CC InterPro; IPR000719; Prot kinase.
 CC InterPro; IPR008271; Ser_thr_pkin_AS.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot kinase; 1.
 CC SMART; SMO0320; S_TK; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
 KW ATP-binding; Phosphorylation; Acetylation; 3D-structure.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT DOMAIN 296 551 PROTEIN_KINASE.
 FT DOMAIN 272 279 POLY-PRO.
 FT DOMAIN 585 597 CALMODULIN-BINDING.
 FT NP_BIND 302 310 ATP (BY SIMILARITY).
 FT BINDING 325 325 ATP (BY SIMILARITY).
 FT ACT_SITE 417 417 BY SIMILARITY.
 FT CONFLICT 335 335 K -> KK (IN REF. 2 AND 3).
 FT HELIX 581 596
 FT TURN 596
 SQ SEQUENCE 607 AA; 65337 MW; 6E677641751A04C8 CRC64;
 Query Match 5.3%; Score 465.5; DB 1; Length 607;
 Best Local Similarity 29.2%; Pred. No. 9.2e-10;
 Matches 131; Conservative 86; Mismatches 174; Indels 57; Gaps 13;
 QY 43 LHSLEPALPGPPSMQVITIEDVQAQTGTAQFAAIEGDP-----QPSVWYKDSVQLVD 96
 DB 156 LHS-----PSCPAAIIAISTEKLPAQKPLSEASELIFEGVPATPGPTPPGPAKEGGVDLLA 210
 QY 97 STRLSQOQEG-----TYSILVLRHVASKDAGVYVTCIAQNTGGQVLC--- 137
 DB 211 E---SQEAGEKAPGQADQAKVGDTSRGIEFQAVPSE-----RPRVEVQALCLPAR 260
 QY 138 KAEILLVGGDNEPDSKQSHR-----RKLHGFEYEV--KEEIGRGVFGVKRVQHKGNKI 189
 DB 261 EEDCFQILDCCPPPPAPPHRIVELATGNVSSEFSNMSKEALGGKGFAGVCTCTEKSTGL 320
 QY 190 LCAAKFTPLRS-RTAQAQVREDDLAALSHPLVTLGDDFETRKTLILILECSSEELLD 248
 DB 321 KLAAKVIKQTQPKQKEMWLEEVNQLNHRNLIQLYAAIEFPHEIVLFMEYIEGGLFE 380
 QY 249 RLVRKGV-VTEAEVKYIQQQLVEGLHSHGVLDHKPKSNILMVHPAREDIKICDFGF 307
 DB 381 RIVDEYHLTEVDTMVFVRQICDGLFMEKRVLHDLKPFENILCVNTTGHVVKIIDEFL 440
 QY 308 AQNITPAELQFQSGSPFVPEIIQQNPVSEASDIWAMGVISYLSLTCSSPFADESRA 367
 DB 441 ARYNPNEKLKYNFGTFPELSEVNVYDQISKTDMSLGVITYMLLSGLSPFLDDDE 500
 QY 368 TLNVLEGRVSWSSPMAHLSEDAKDFIKATIQAPQAPRPSAAQCLSHFWFKLSPASEA 427
 DB 501 TLNVLSGNWYFDEETFEAVSDEAKDFVSNLIVKEQGARMASAAQCLAHWP-LNNL-ASKA 558
 QY 428 HFINTK-----QLKPLLRSRWQRISMS 450
 DB 559 KRCNRRLKSQLLKXVLMKRWKKNFIA 586
 RESULT 15
 ULK1 HUMAN STANDARD; PRT; 1050 AA.
 ID ULK1 HUMAN
 AC O75385;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine/threonine-protein kinase ULK1 (EC 2.7.1.-) (Unc-51-like kinase
 1).
 DE ULK1.
 GN ULK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98360094; PubMed=9693035;
 RA Kuroyanagi H., Yan J., Seki N., Yamanouchi Y., Suzuki Y.-I.,
 RA Takano T., Muramatsu M.-A., Shirasawa T.;
 RT "Human ULK1, a novel serine/threonine kinase related to UNC-51 kinase

Search completed: April 23, 2004, 15:09:29
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2004, 15:04:47 ; Search time 63 Seconds
(without alignments)

8338.695 Million cell updates/sec

Title: US-10-697-263-2

Perfect score: 8740
Sequence: 1 MGCCRLGCGCSVAHSVSG.....RNRKERRALLYKRNLAQVR 1665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_plant.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6898	78.9	1319	4 Q9HCD3	Q9hcd3 homo sapien
2	1913	21.9	2242	4 Q9P2P9	Q9p2p9 homo sapien
3	1864	21.3	3262	11 Q9EQJ5	Q9eqj5 mus musculus
4	1547.5	17.7	341	11 Q8QUP9	Q8qup9 mus musculus
5	858.5	10.3	1171	11 Q8QTF7	Q8qtf7 mus musculus
6	824	9.4	8081	5 Q7Z1Z0	Q7z1z0 caenorhabdi
7	819	9.4	799	11 Q8BZF4	Q8bzfa mus musculus
8	701.5	8.0	1393	5 Q7Z1I9	Q7z1i9 caenorhabdi
9	695.5	8.0	1226	5 Q9WID6	Q9wid6 drosophila
10	695.5	8.0	3197	5 Q9WID5	Q9wid5 drosophila
11	580.5	6.6	1042	11 Q8QW23	Q8qw23 mus musculus
12	552	6.3	1721	5 Q961U1	Q961u1 drosophila
13	552	6.3	8943	5 Q9VAF7	Q9vaf7 drosophila
14	551.5	6.3	8625	5 Q8EGD6	Q8egd6 procambarus
15	541	6.2	992	4 Q9COL5	Q9col5 homo sapien
16	541	6.2	1914	4 Q7Z4J0	Q7z4j0 homo sapien

17	538	6.2	929	5	O01651	O01651 drosophila
18	536.5	6.1	1289	4	Q9V2A5	Q9v2a5 homo sapien
19	536	6.1	641	6	Q9B3E9	Q9b3e9 macaca fasc
20	534	6.1	2959	11	Q9JIFI	Q9jifi rattus norv
21	531	6.1	6558	5	O76281	O76281 drosophila
22	529	6.1	529	5	Q8MLD7	Q8mld7 drosophila
23	529	6.1	1506	5	Q9V7G6	Q9v7g6 drosophila
24	529	6.1	4463	5	Q8MLD8	Q8mld8 drosophila
25	529	6.1	9270	5	Q8MLD9	Q8mld9 drosophila
26	528	6.0	611	11	Q91XS9	Q91xs9 cavia porce
27	526	6.0	4736	5	Q7VT99	Q7vt99 mytilus gal
28	523.5	6.0	2780	5	Q8MNS0	Q8mns0 caenorhabdi
29	523.5	6.0	2808	5	Q8MNS1	Q8mns1 caenorhabdi
30	523.5	6.0	18519	5	Q8ISF6	Q8isf6 caenorhabdi
31	519.5	5.9	1129	11	Q8QOUX0	Q8qoux0 mus musculu
32	508	5.8	1950	11	Q8QYNS8	Q8qyns8 mus musculu
33	507	5.8	18534	5	Q8ISF7	Q8isf7 caenorhabdi
34	506.5	5.8	878	5	Q9GV22	Q9gv22 mytilus gal
35	504.5	5.8	832	5	O01653	O01653 drosophila
36	503	5.8	2693	5	Q8ISF3	Q8isf3 caenorhabdi
37	501.5	5.7	1031	11	Q8QYN7	Q8qyn7 mus musculu
38	501	5.7	907	13	Q988E0	Q988e0 carassius a
39	500	5.7	785	5	Q8MLD6	Q8mld6 drosophila
40	498	5.7	6831	5	Q23550	Q23550 caenorhabdi
41	498	5.7	7158	5	Q23551	Q23551 caenorhabdi
42	497.5	5.7	732	5	Q9GTV1	Q9gtv1 drosophila
43	496	5.7	786	5	Q9GV79	Q9gv79 drosophila
44	492	5.6	724	11	Q8BWD1	Q8bwd1 mus musculu
45	489	5.6	448	11	O54784	O54784 mus musculu

ALIGNMENTS

RESULT 1

Q9HCD3 PRELIMINARY; PRT; 1319 AA.
ID Q9HCD3
AC Q9HCD3;
DT 01-MAR-2001 (TREMSLrel. 16, Created)
DT 01-MAR-2001 (TREMSLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMSLrel. 25, Last annotation update)
DE Hypothetical protein KIAA1639 (fragment).
GN KIAA1639.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450693; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046859; BAB13465.1; -.
DR HSP; P56276; 1TLK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003931; FN.II.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG.c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008266; Tyr_kinase_As.
DR Pfam; PF00047; IG.1.
DR Pfam; PF00069; pkinase; 2.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IG2; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.

DR PROSITE; PS00109; PROTEIN KINASE TYR: 1.
 KW Hypothetical protein; ATP-binding; Immunoglobulin domain; Transferase.
 FT NON_TER 1
 SQ SEQUENCE 1319 AA; 142049 MW; FBA43AE17204EF48 CRC64;
 Query Match 78.9%; Score 6898; DB 4; Length 1319;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1318; Conservative 1; Mismatches 0; Indels 2; Gaps 1;
 QY 345 AMGVISYLSITCSPPFAGESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLORAPQ 404
 DB 1 AMGVISYLSITCSPPFAGESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLORAPQ 60
 QY 405 ARPSAAOCLSHPWFLKSWPAEAEHINTKQLFLLARSQRMSYKSLVMSRISPELL 464
 DB 61 ARPSAAOCLSHPWFLKSWPAEAEHINTKQLFLLARSQRMSYKSLVMSRISPELL 120
 QY 465 RGPDPSPGLVARHLCDRTGGSSSSSSSDNELAPFAFAKSLPPSPVTHSPILHPRGFLR 524
 DB 121 RGPDPSPGLVARHLCDRTGGSSSSSSSDNELAPFAFAKSLPPSPVTHSPILHPRGFLR 180
 QY 525 PSASLPEAEASERSTAP 584
 DB 181 PSASLPEAEASERSTAP 240
 QY 585 APGRRHPARRRHLLKGGYIAGALPGLREPLMEHRVLEEAAREEQATLLAKAPSPETAL 644
 DB 241 APGRRHPARRRHLLKGGYIAGALPGLREPLMEHRVLEEAAREEQATLLAKAPSPETAL 300
 QY 645 RLPASGTHLACHSHLSLHSDSPSPSPSSSEACGAQLPSAPSGCAPIRDMGHPQGSQKL 704
 DB 301 RLPASGTHLACHSHLSLHSDSPSPSPSSSEACGAQLPSAPSGCAPIRDMGHPQGSQKL 360
 QY 705 PSTGHPGTAQPERSPSPGQAPFCHPKQGSAPQEGCGPHPAVAPCPGPPSPGSCCK 764
 DB 361 PSTGHPGTAQPERSPSPGQAPFCHPKQGSAPQEGCGPHPAVAPCPGPPSPGSCCK 420
 QY 765 EAPLYPSFFFLGQAPAPAKAP 824
 DB 421 EAPLYPSFFFLGQAPAPAKAP 480
 QY 825 VSSLVGSSQVGTETPGSLDAEGMTQEAEDLSDTPTLQRPQEQVTKRKSGLGRGGYAG 884
 DB 481 VSSLVGSSQVGTETPGSLDAEGMTQEAEDLSDTPTLQRPQEQATWRKFSGLGRGGYAG 540
 QY 885 VAGYGTFAFGDAGMLGQGMWARIWAVSQSEEEHQAARASQSEEQEAEAEPLP 944
 DB 541 VAGYGTFAFGDAGMLGQGMWARIWAVSQSEEEHQAARASQSEEQEAEAEPLP 600
 QY 945 QVSARVPVEVGRAPTRSPPTPMDIGQVSLVQIRDLSDGDAEADTISLDSISEVDPAYL 1004
 DB 601 QVSARVPVEVGRAPTRSPPTPMDIGQVSLVQIRDLSDGDAEADTISLDSISEVDPAYL 660
 QY 1005 NLSLDYDIKYLPEEFMIFRKVPKSAQPEPPSPMAEELAEPEPTWMPGELGPHAGLEI 1064
 DB 661 NLSLDYDIKYLPEEFMIFRKVPKSAQPEPPSPMAEELAEPEPTWMPGELGPHAGLEI 720
 QY 1065 TESEEDVDALLAAVAVGRKRWSSPSRSLFHPGRLHPLDEPAELGLRERKASVEHISR 1124
 DB 721 TESEEDVDALLAAVAVGRKRWSSPSRSLFHPGRLHPLDEPAELGLRERKASVEHISR 780
 QY 1125 ILKGRPEGLKSGPRKPKGLASFLRSLGKSWDRAPTFLRELSDTIVLGGSVTLACQVS 1184
 DB 781 ILKGRPEGLKSGPRKPKGLASFLRSLGKSWDRAPTFLRELSDTIVLGGSVTLACQVS 840
 QY 1185 AQFAAATWSKQAPLLESSRVLISATLKNFOLLILVVAEDLGVTTCVSNALGTVT 1244
 DB 841 AQFAAATWSKQAPLLESSRVLISATLKNFOLLILVVAEDLGVTTCVSNALGTVT 900
 QY 1245 TGVLRKAERPSPPCDIGEVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTTLASDIF 1304
 DB 901 TGVLRKAERPSPPCDIGEVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTTLASDIF 960

QY 1305 DCCYLTSKLSRGTYTFTACVSKAGMGPYSSPSEQVLLGSPSHLASEESQGRSAQPLP 1364
 DB 961 DCCYLTSKLSRGTYTFTACVSKAGMGPYSSPSEQVLLGSPSHLASEESQGRSAQPLP 1020
 QY 1365 STKTFAPOTQORGFPSVVRQWCKASGRALAAKIIPVHPKDKTAVLREYALKGLRHPH 1424
 DB 1021 STKTFAPOTQORGFPSVVRQWCKASGRALAAKIIPVHPKDKTAVLREYALKGLRHPH 1080
 QY 1425 LAQLHAAVLSPRHLVLIILELCSGPPELLPCLAEASYSSEVKDYLMQWLSATQVILHNOHI 1484
 DB 1081 LAQLHAAVLSPRHLVLIILELCSGPPELLPCLAEASYSSEVKDYLMQWLSATQVILHNOHI 1140
 QY 1485 LHLDIRSENMIITEYNLLKVVLDLGNQASLSQKVLPSDKPKDYLETMAPELLEGQGVAPQ 1544
 DB 1141 LHLDIRSENMIITEYNLLKVVLDLGNQASLSQKVLPSDKPKDYLETMAPELLEGQGVAPQ 1198
 QY 1545 TDIWAIGVTAFTMLSAEYFVSSEGARDLQRLKGLVRLSRCYAGLSGGAFAFURSTLCA 1604
 DB 1199 TDIWAIGVTAFTMLSAEYFVSSEGARDLQRLKGLVRLSRCYAGLSGGAFAFURSTLCA 1258
 QY 1605 QWGPFCASSCLOCWLTTEGPACSRPAPVTPPTARLEVFVRNREKRALLVKRNLQAV 1664
 DB 1259 QWGPFCASSCLOCWLTTEGPACSRPAPVTPPTARLEVFVRNREKRALLVKRNLQAV 1318
 QY 1665 R 1665
 DB 1319 R 1319
 RESULT 2
 Q9P2P9 PRELIMINARY; PRT; 2242 AA.
 ID AC Q9P2P9
 DT 01-OCT-2000 (TREMELrel. 15, Created)
 DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical protein KIAA1297 (fragment).
 GN KIAA1297.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K., Hirotsawa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes.XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AB037718; BAA92535.1; -.
 DR HSP; P56276; ITLK.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004574; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR00719; Prot_kinase.
 DR InterPro; IPR008271; Ser_Thr_kin_AS.
 DR Pfam; PF00041; fn3; 2.
 DR Pfam; PF00047; ig; 5.
 DR Pfam; PF00069; pkinase; 2.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00408; IGC2; 3.
 DR PROSITE; PS00835; IG_LIKE; 5.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 2.

DR	PROSITE; PS00108; PROTEIN KINASE ST; 2.	
KW	Hypothetical protein; ATP-binding; Immunoglobulin domain; Kinase;	
KW	Serine/threonine-protein kinase; Transferase.	
PT	NON_TER	1
FT	NON_TER	2242
SQ	SEQUENCE	2242 AA; 243491 MW; 497946BS26A3561A CRC64;
	Query Match	21.9%; Score 1913; DB 4; Length 2242;
	Best Local Similarity	29.3%; Pred. No. 5.1e-114;
	Matches	555; Conservative 229; Mismatches 611; Indels 500; Gaps 51;
QY	54 PSMQVTTEDVQAQGGTAQFAEIIIEGDPQPSVTKYKDSVOLVSTRLSQOQGGTYVSLV	113
Db	504 PRFSIMEDVVGAGETARFAVVVEGKPLDPMYKDEVLLTSSHVSVFYBENECSLVV	563
QY	114 RHVASDAGVYTCIAQNTGGCVLCKABELLVLGDN-----EPDSEKQSHR-RKLHSFYEVK	168
Db	564 LSTGAQDGVVTCIAQNLAGEVSKAEALVHSAQTAAMEVEGVGEDSDHGRRLSDFYDIH	623
QY	169 EIGRGVFGFKVRQHKNGKILCAKEPIPLRSRTRAQAYRERDIIAALSHPLVTGLLDOF	228
Db	624 QEIRGAFSYLRRIVERSSGLFEPAKFIPOAKPKASRREARLLARLOHDCVLYFHEAF	693
QY	229 ETRKTLIIELCSEBELLRLVRKGVVTEARVKVITQQLVEGLHSHGVHLHDIKPS	288
Db	694 ERRGLVITVTELC-TEELLERIRARKPTVCSEIRAYMRQVLEGIHYLHSHVHLHDVKE	742
QY	289 NILMVHPA--REDIKICDFGAQNIITAELOFQYQSGSPFVSEPIITQQPVSEADINAM	346
Db	743 NLLVMDGAAGQQVRIICDFGNAOELTEGEPOYCYGTPEFAVEPIYNQSPVSGVTDINEV	802
QY	347 GVISYLSITCSPFAGSDRATLNVLEGRVSWSSPMAAHLSDAKDF-IKATLQRAPOA	405
Db	803 GWAFCLCTGISPFVGENDRITLNIIRNVNVAFEETFLSLSEARGFLIKVLVQ--DRL	860
QY	406 RPSAAOCLSHWFLKSPABEAHFINTKQLKFLARQORSLSMYSKLSVIMRSPELLR	465
Db	861 RPTAETLEHFWPKTQAKGAE---VSTDHLKFLSRRRQORSQISYKCHLVLRPIPELLR	917
QY	466 GPPDSPSLGVARHLCDRTGGSSSSSSSSDNEL-----APPARAK-SLPPSPVTH	513
Db	918 APPRVVVTWERR--PPSGGLSSSSDSEEELEELPSVPRPLQPESGSRVSLDIPTED	976
QY	514 SPLHPGFLRPSASLPEEABERSTEAPAPASPEGAPPAQCVPRHSHVIR-----	568
Db	977 EALGTPTGAATPMQWQEQRAFSQOEAPSPALFSPQEPAA--GASPRRGELRGSSA	1035
QY	569 -----SLFYHQAG--ESPEHGALPG-----SRHFAARRHLIK	600
Db	1036 ESALPRAGPRELGRGLHKAASVELPCRRSPGCGATRLARGGLGEGRYACQLQALLR	1095
QY	601 GGYTAGLPLGLRPLMEH-----RVLEEEAAREEQATL-----LAKAPSFETALR	645
Db	1096 GGPEDGVKVLGRPLLESGLGRARDPRMAAASSEAPHQPLENRGLQKSSSFSGEA	1155
QY	646 LPASGTHLAPCHSHLSHSDSPSTPR-----PSSEACGEAQLRPSAGGAPFIRDMGHPOGS	701
Db	1156 EP-RGRHRRAGAPLEIFVARLGARRLOESFSLALSEAQ--FSSPA-----RPSAP	1203
QY	702 KQLSTGCHGCTAQRPSRPSDPGWOPAPPCHPKQGSAPQEGCSHPAVAPCPGSPFP	761
Db	1204 X--PST---PKSAEPSATTPSDAPQPPAP--QPAQDKAPEPRPEPVASKPAP-----	1252
QY	762 SKC--EAPLVPSPFLQ-----PQAPPAPAK-----ASPPLDKMGF	797
Db	1253 ALQTLALPLTFYAGIIQSLQLSGHAQGPSQGPAPPESEPKPHAAVFARVASP-----	1305
QY	798 GDISLPGRKPGKCSGPGSASQASSQVSLRVSSQVGTGPGSPSLDAEGWTOEADLSD	857
Db	1306 -----PGAPEKRVPSAGGPVPLAEKARVPT-----VPPRGSSLSLSSENLESEAVFE	1353
QY	858 STPTLQRPQEQVTRKKSLSGRGGYAGVAGYGTGTFAGDAGGMLGQGPWWAXIANAVSQS	917

RESULT 3
Q9EQJ5 ID Q9EQJ5 PRELIMINARY; FRT; 3262 AA.
AC Q9EQJ5, (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)

DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)	QY	406	RPSAAQCLSHPWFLKSMPEAEAHFINTKQLKFLLRGRWORSLSYKSIILVMRIPELLR	465
DE	Striated muscle-specific serine/threonine protein kinase.	Db	1847	RTABETLEHPWFKTEAKGAE---VSTDLKLFSLRRWQSQISKCHLVLRIPELLR	1903
GN	AP031 OR SPEG.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	QY	466	GPDPSPSLGVARHLCDRTGGSSSSSSSSDNL	513
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Db	1904	APPERVWAMPRRQ--PPSGGLSSSSDSEEELELPSVPRPLQPEFGSGRSVLTIDTIED	1962
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.	QY	514	SPLLHPRGFLRPSASLPEAEASERSTEAPAPPA	558
RC	STRAIN=BALB/c;	Db	1963	EALGTPEAGAAATPMQWQERTPSKQAPSPALPQSQSSPDGPRRPELRGSSAE	2022
RX	MEDLINE=20538441; PubMed=10973969;				
RA	Hsieh C.M., Fukumoto S., Layne M.D., Maemura K., Charles H., Patel A.,	QY	559	GCVPR---HSVIRSLFYHQAGSEPHGALAPG	601
RT	"Striated Muscle Preferentially Expressed Genes alpha and beta Are Two	Db	2023	SALPRVGRREPRSLHKAASVELPQRSPSPGATRLTRGGLGEYACRLQRLRLRG	2082
RI	Serine/Threonine Protein Kinases Derived from the Same Gene as the	QY	602	GYIAGALPGLREPLMEHREVEAREEQATLLAKAPSFETALRLPASGTHLAPGHSHL	661
RL	Aortic Preferentially Expressed Gene-1."	Db	2083	GPEDGKVSGLRGL---LESGLGR---ARDPMARAAASSEAAHPHQPPEPESRGL	2130
RJ	J. Biol. Chem. 275:36966-36973(2000).				
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.	QY	662	EHDS-----PSTPRPSSACGAEQRLPSAPSGGPIRDMGHPOG--SKOLSTGGHPG---	712
DR	EMBL; AF215896; AAC34791.1; -	Db	2131	QKSSSFQGEAEPRGRHRRAGAPLEIFVRLGARRLQESLSLSALSTQPPSPARPVFK	2190
DR	HSSP; P56276; 1TLK.	QY	713	---TAQPE-----RPPSPDPGWQAPPHCHPKQSGAPQEGCSP--HPAVAPCPGSPSPFP	761
DR	GO; GO:0005224; F:ATP binding; IEA.	Db	2191	LSITKSPSPSAVTSRSDSPQPEPQVPEKVPPEKPEPVRAAKPAQPLALQMTQPTLPY	2250
DR	GO; GO:0016740; F:transferase activity; IEA.	QY	762	SCKEAPLVSPSPFLGQCPQAPAPAKASPPLDSKMGPGDISLPGRPKPGCPSSPQASQAS	821
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.	Db	2251	AQIMQSLQSLSTL--SPQDP-----AVPSPKPAHFAVARFARVAPPPGV-----SEKVP	2299
DR	InterPro; IPR001961; FN III.	QY	822	SSQVSLRAGVSGVQGT---RPGSLDABGTQAEADLSDSTPTLQRPQEQVTKRKFSLGG	878
DR	InterPro; IPR007110; IG-like.	Db	2300	SARTPFLAEKARVTPVPRFGSLSGSLENLEAEVFEA--KFRGRSFLSRGLL---	2355
DR	InterPro; IPR003598; IG c2.	QY	879	RGYAGVAGYGTFAFGDGAGMLGQPMWARIANAVSQSEEEQEAARASQEEQEAR	938
DR	InterPro; IPR000719; Prot kinase.	Db	2356	-----LSRSRSEERGPFRGAEDDGIYRPS	2380
DR	InterPro; IPR008271; Ser_Thr_pkin_AS.	QY	939	ABSPLPQVSARVPVGVGRAPTRSPPTPWNEDIGQVSLVQIRDLGSDAEADTISLISE	998
DR	Pfam; PF00047; Ig; 8.	Db	2381	ACTPL-----ELVRRPERSR-----VQDLRVAGPGLVRRLSLSLSQ	2418
DR	Pfam; PF00069; pkinase; 2.	QY	999	VDPAVLNLDLYDIKYLPFEFMFRKVPKSAQPEPPSPMAEEELAEPEPTWMPGLGP	1058
DR	ProDom; PD000001; Prot kinase; 2.	Db	2419	K-----LRTPPGQR-----HPAESRSQD--	2439
DR	SMART; SMC0060; FN3; 2.	QY	1059	HAGLEITESEVDALLAEAVGRKRWSPSRSLFH-----FPGHLDLDE	1105
DR	SMART; SMC0408; IGC2; 5.	Db	2440	-----ESSGGSSARASPVLAVERRLSTLERLSRLQSSSEDSGSGAGKSTPL--	2490
DR	PROSITE; PS00815; IG LIKE; 8.	QY	1106	PAELGLRERVKASVEHISRI-----LKGPRGLEKEGPD-----	1139
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	Db	2491	---FGLRLRATSEGESLRLGVPHNQLGSGTATTTPSAESLGSASCTSGSSAPGESR	2547
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 2.	QY	1140	-----RKPGLASFLSLG-----KSWDRAFTFLLELSDETVVLGQSVTLA	1180
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 2.	Db	2548	HRWGLSLRLKQGLSQPNLSSVQEDLGHQVYPSSEDPPVFFHKLKQDQVLLGEAATLL	2607
KW	ATP-binding; Immunoglobulin domain; Kinase;	QY	1181	CQVSAQPAQAQATWSKDGAPLESRSRSLISATLKNFQLLTILVVVAEDLVYTCVSNALG	1240
SW	Serine/threonine-protein kinase; Transferase.	Db	2608	CLPAACAPRISMMKQKQSLSESPVIVSCDKQQLLSIPRACKRHAGLYECSATNVLG	2667
SQ	SEQUENCE 3262 AA; 354280 MW; C490471FC26E0255 CRC64;	QY	1241	TVTTTGVLRKARPPSSPCPDIGEVADGVLLVHKPVESGPTVYVQCSLEGGS-WTTL	1299
		Db	2668	SITSCCTAVARIPOKLAPPVQTYHTALVWVGDPGRAPCTYTLERRVDGESVWHVP	2727
		QY	1300	ASDIFDCCLYTKSLRGGTYTFRACVSKAGMGYPSPSEQLVGG-----	1345
		Db	2728	SSGIPDCYVNTVQLPVGVTVFRVACSNRACQGPSPSEKVFIRGTGTPSPAAPAAAPRD	2787
		QY	1346	-----	1345

Query Match	21.3%; Score 1864; DB 11; Length 3262;
Best Local Similarity	28.3%; Pred. No. 1.3e-110;
Matches 540; Conservative 238; Mismatches 688; Indels 440; Gaps 42;	
QY	54 PSMQVETEDVQAGTGTAFQAEALIEGDQPPSVTWKDSVQLVDSTRLSQQSGTYSVLV 113
Db	1490 PFESEMEDVEVGPGETAPAVVVEGKPLDWMYKDEVLAEASHVSVFVEECSVLV 1549
QY	114 RHVASKDAGVYVTCIAQNTGGOVLCKAELLVLGDN---EPDSEKQSHR-RKLHSHFYBVK 168
Db	1550 LSAGSQGGVYVTCIARNLAGEVSKAELSVLSAQTNMEVGVGEDEHGRRLSDYDIH 1609
QY	169 BEIGRGVGFVKRVQHKGNKILCAAKPIFLRSTRQAQAVREDIILAAHSHPLVTGLDQF 228
Db	1610 QEIGRGAFSLRRVVRSSGLSEFAKFIQSPKAPKASARAEARLLARLQHGCVLYFHEAF 1669
QY	229 ETRKTLILILECCSELDDLRYKGVWTEAEVKVVIQQLVEGLHYLHSHGLHLDIKPS 288
Db	1670 ERRRLVIVTSLC-TEELLERMARKPTVCESETRVTRQVLEGI CYLQSHVHLVDKPE 1728
QY	289 NILMWHFA--REDIKICDPGFAQNTIPAELOFSQSGSPFVSPRIIQNPVSEASDIAM 346
Db	1729 NLLVMDGAGGEBQVRCIDFNGAQLTFFGFCQVGTGTFEFAVPEIWNQSPVSGVTDIPV 1788
QY	347 GVISYLSITCSSPFAESDRATLLNVLGRVSKSSPMAAHLSEDAKDF--IKATLQRAQPA 405
Db	1789 GVAFLCLTGISFPVGENDRTLLMNIRYNVAFEETFLSLSREARGLIKVLVQ--DRL 1846

```
Db 2788 APVTSGETRAPPPDPTSLAPALAPPASQASTLSPTSSMSANQALSSLKAVGPPPAT 2847
QY 1346 -----PSHLASEEE-----SQRSAPLSTKTFAF-- 1371
Db 2848 PPKHRLATQAAEBSPPSIVVTPSEPRSFVDPDTGTLTPSSPGCVKPAFSSISLYNVT 2907
QY 1372 -----QTOIQORGFVVRQCKEKA 1390
Db 2908 SFVSAPPAPAPAPPEPPPTKVTVRSLSPAKEVVSSTPTSTLRQGLFLENPTFSWRR 2967
QY 1391 SGRALA-----AKIIPYHPKDTAVLREYKALGRLHPLAQLAAALYLS 1435
Db 2968 PGALALCGHAGMLRAERLSPRFVYAAEGRRVLQEVVLRTLHBERLSLHEAVITP 3027
QY 1436 RHVLVILELCSGPPELLPCLAEASVSEVKYLQWLSATQYLHNLHLDLRSNMI 1495
Db 3028 RYLVLVAESCGNRELCCGSDPRYSDDVATYVVQLQGLDYLGHVHLHLDIKPNLL 3087
QY 1496 ITEYNLLKVVLDGNAQSLSQEKLPSDFKDYLETWAPLELGGQAVPQTDIWAIGVTAF 1555
Db 3088 LAADNALKIVDFGSAQPYNFQALKPLGHRTGTLEFMAMPVKGDPIGSAITDINGAGVLY 3147
QY 1556 IMLSABYVSSSECDRLQRLKGLVRLSRVAGLSGGAFAFLSTLCAQWGRPCASSC 1615
Db 3148 IMLSGSPYEPDPOETEARIIVGGRFDAFQLYPNTSQSATFLKRLVSVHFWSRPSLQDC 3207
QY 1616 LQCPMLTEBGPACSRPAPVPTFTABLRVFNREKRRALLYKRNLL 1661
Db 3208 LAHPQLDAYLMKLRQTITFTTNELKSPGQRERRAEATRKV 3253

RESULT 4
ID Q80UP9 PRELIMINARY; PRT; 341 AA.
AC Q80UP9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LOC216790 protein (Fragment).
GN LOC216790.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Breast tumor;
RA Strausberg R.,
```

```
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046431; AAH46431.1; -.
DR GO; GO:0005524; F.ATP binding; IEA.
DR GO; GO:0004674; P.protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P.protein-tyrosine kinase activity; IEA.
DR GO; GO:0000468; P.protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
FT NON_TER 1
SQ SEQUENCE 341 AA; 38078 MW; 63BDB2E3FFD71914 CRC64;

Query Match 17.7%; Score 1547.5; DB 11; Length 341;
Best Local Similarity 87.7%; Pred. No. 1-2e-91;
Matches 299; Conservative 16; Mismatches 25; Indels 1; Gaps 1;

QY 1326 VSKAGNGYSPSEVOVLGGPSHLASEEE-SQRSAPLSTKTFAFQIQORGFVVR 1384
Db 1 VSKAGNGYSPSEVOVLGGPSHLASEEE-SQRSAPLSTKTFAFQIQORGFVVR 60

QY 1385 QCKEKASGRALAAKTIIPYHPKDTAVLREYKALGRLHPLAQLAAALYLSPRHLVLI 1444
Db 61 QCKEKASGRALAAKTIIPYHPKDTAVLREYKALGRLHPLAQLAAALYLSPRHLVLI 120

QY 1445 CSGPELLPCLAEASVSEVKYLQWLSATQYLHNLHLDLRSNMIITYNLLKV 1504
Db 121 CSGPELLPCLAEASVSEVKYLQWLSATQYLHNLHLDLRSNMIITYNLLKV 180

QY 1505 VDLGNAQSLSQEKLPSDFKDYLETWAPLELGGQAVPQTDIWAIGVTAFIMLSAEYV 1564
Db 181 IDLGNAQSLDQEKVPAPNFQDYLETWAPLELGGQAVPQTDIWAIGVTAFIMLSGEY 240

QY 1565 SSEGARDLQRLKGLVRLSRVAGLSGGAFAFLSTLCAQWGRPCASSCLOCQPWTEE 1624
Db 241 SSEGARDLQRLKGLVRLSRVAGLSGGAFAFLSTLCAQWGRPCASSCLOCQPWTEE 300

QY 1625 GPACSRPAPVPTFTABLRVFNREKRRALLYKRNLLAQVR 1665
Db 301 GPTGSRPTPTFTVRLRAFAVREKRRALLYKRNLLAQVR 341

RESULT 5
ID Q80TF7 PRELIMINARY; PRT; 1171 AA.
AC Q80TF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MKIAA1297 protein (Fragment).
GN MKIAA1297.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
II. The complete nucleotide sequences of 400 mouse KIAA-homologous
cDNAs identified by screening of terminal sequences of cDNA clones
randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
RN [2]
DB EMBL; AK122488; BAC65770.1; -.
```

DR	GO:	GO:0005524; F-ATP binding; IEA.
DR	GO:	GO:0004674; F-protein serine/threonine kinase activity; IEA.
DR	GO:	GO:0004713; F-protein-tyrosine kinase activity; IEA.
DR	GO:	GO:0006468; P-protein amino acid phosphorylation; IEA.
DR	InterPro:	IPIR003961; FN III.
DR	InterPro:	IPIR008957; FN_III-like.
DR	InterPro:	IPIR003599; IG.
DR	InterPro:	IPIR007110; IG-like.
DR	InterPro:	IPIR003598; IG_c2.
DR	InterPro:	IPIR000719; Prot Kinase.
DR	InterPro:	IPIR002290; Ser Thr_kinase.
DR	InterPro:	IPIR008271; Ser thr_kin_AS.
DR	InterPro:	IPIR001245; Ty*_kinase.
DR	Pfam:	PF00041; fn3; 1.
DR	Pfam:	PF00069; pkinase; 1.
DR	ProDom:	PD000001; Prot_kinase; 1.
DR	SMART:	SM00409; IG; 1.
DR	SMART:	SM00408; IGC2; 1.
DR	SMART:	SM00220; S_TKc; 1.
DR	SMART:	SM00219; Ty_KC; 1.
DR	PROSITE:	PS50835; IG_LIKE; 1.
DR	PROSITE:	PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE:	PS00108; PROTEIN_KINASE_ST; 1.
FT	NON_TER	1
SQ	SEQUENCE	1171 AA; 126323 MW; BAE413056089AE57 CRC64;

Query Match	10.3%; Score 898.5; DB 11; Length 1171;
Best Local Similarity	24.5%; Pred. No. 5.6e-49;
Matches 318; Conservative 130; Mismatches 468; Indels 381; Gaps 29;	

QY	611	LREPLMEHRVLEEEAAAREEQATLLAKAPSFETALRLPASCTHLPAGHSLSLHDS-----665
DB	1	LRGPLL-----LESIGGR-----ARDPRVARAAASSEAAHPHOPPESEGLCKSSFSFQG 48
QY	666	PSTPRSPSEACGAQLPSAPSGCAPIRDNGHPQG-SKLPTSGCHPG-----TAQPE- 717
DB	49	EAEPPGRHRRAGALEIIVAEALGARRIQESPSLSALSETQPPSPARPSVPKILSIKTSPEP 108
QY	718	----RPSPDPSGWGPAPFCHFGKGSAPQEGCSFP--HPAVACPCCGPSPPGSCKEARLVLP 770
DB	109	SAVTSRDSPQPPEPCQVEPKYPEKPEPVZAAKPAQPLALQMTPQLPTPAQIMQSILQL 168
QY	771	SPFFLGQQAPPAPAKASPPLDXMGCGDISLGRPKPGPCSPGSAQSSSQVSILRV 830
DB	169	SSFTVL-SFQDP-----AVPSEPPEPHAAVFARVASPPPVG-----SEKRVPSARTPPVLA 217
QY	831	GSSQVGT---BPGSLDAEGWTQEAEDLSDTPTLQRPOQTVMRKFSILGGRGGVAGVAG 887
DB	218	EXARVTVPPRPGSLSGSITENLESEAVFEA--KFKRSRESPLSLGLRL-----264
QY	888	YGTAFPGSDAGMLGQGMWARIATAWYSQSEEEEEQEERAAESQEQEARAESPFLPQVS 947
DB	265	-----LSRSRSSEERGPFPGAEDDGIYRPSAGTPL-----294
QY	948	ARPEVEVGRATRSPPEPTWEDIGQVSLVOIRDLSDAEADATTLSLISEVDPAYNLNLS 1007
DB	295	-----ELVRPERERS-----VODLRVAGEPGLVRRLSLSLSOK-----328
QY	1008	DLYDIKYLPFBFMIFRKVPKSAQEPPEPSPMAEBELAEFPETWPMPGSLGHAGLEITEE 1067
DB	329	-----LRNTPPQQR-----HPAWESRSGDG-----ES 350
QY	1068	SEDVDALLAAAVGRKKWKNSPERSLFH-----PPGRHLPLDEPSELGLURER 1114
DB	351	SGGSSAPASPVLA VRRLSLSTTLRSSLRQSSGEDSGGASGRSTPL-----FGRLRR 405
QY	1115	VKASVEHISRL-----LKGRPEGLEKEGP-----R 1140
DB	406	ATSEGESLRLRLGVPHNQIGSQTGATTPTSABESLGSASGTSGSSAFGESRSRHWGLSLRL 465
QY	1141	KPGGLASFLISGL-----KWDRAPTFLRELSDDETIVLIGQSVTLACQVSAQAPAA 1189
DB	466	KDKGLSQPNLSSSVQEDLGHQVYPSSDFPPFHIXLKQOVLLGEAATLLCLPAACFAP 525

RESULT 6

Q7Z120	PRELIMINARY;	PRT;	8081 AA.
AC	Q7Z120;		
DT	01-OCT-2003 (trEMBLrel. 25, Created)		
DT	01-OCT-2003 (trEMBLrel. 25, Last sequence update)		
DT	01-OCT-2003 (trEMBLrel. 25, Last annotation update)		
DE	Uncoordinated protein 89, isoform B.		
GN	UNC-89.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Peloderinae; Caenorhabditis.		
CC	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RX	MEDLINE=99069613; PubMed=9851916;		
RA	Wilson R.;		
RT	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium."		
RL	Science 282:2012-2016(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	Du Z., Le T.T., Wilson R.;		
RT	"The sequence of C. elegans cosmid C09D1.";		
RL	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.		

RN RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003131; AAP68958.1;
 SQ SEQUENCE 8081 AA; 894245 MW; 67C804953CF62228 CRC64;
 Query Match 9.4%; Score 824; DB 5; Length 8081;
 Best Local Similarity 22.3%; Pred. No. 6.3e-43;
 Matches 405; Conservative 248; Mismatches 661; Indels 502; Gaps 75;
 60 IEDVOAQGGTAQ---FPAIEGPPQSVTWYKDSVOLVDSTRLSQQOEGT-TYSLVLRH 115
 6510 IOLVNTKTFSSAQPMRMDVRVDGPPFELKWKMKWRPIVSSRIKFKVDGGLYLSLIND 6569
 116 VASKDAGVYTLAQNTGGQVLCBAELLVLGGDNPDPSEKQSHR-----RKLHSYEVKKE 170
 6570 PMWRDSGIYSCVAVNDAGATTCTVTVEABEGYNDVLPERRVTIESRRVRELYEISE- 6628
 171 IGRGVGFVQVQKHGKILCAAKFIPLRSRTQAQ-----YRERDILAAAL 216
 6629 -----KDEKL--AEGAPFVKEKATGREFLAQLRPIDDALMGHVDIENSL 6672
 217 SHPLVTGLDOPETKTLILILELCSSELDRLYR-----KGVTEAEVKYI 265
 6673 DHPGIVQHRVLRDEKLALVFD--NANSTIDGLSSLAHPGVEIAEPKGVNRETCSRFFV 6730
 266 QOLVGLHVLHSHGVHLHDIKPSNLMVHPAREDIKIDFGFAQNTITPAELQFQSQYSPSE 325
 6731 RQLLALXKMDLRLAHLDRETFILL--QDDKLKADPQARLLRGLITGKISPE 6787
 326 FVSPRIIQNPVSEASDIWAMGVISYLSLTSSPPAGESDRATLNLVLEGRVSWSSMAA 385
 6788 FVSPVIRSYPLTATDMWSTGLVYLLTGLSPHGDNDNETLANVDSQCFD--SSPL-G 6845
 386 HISEDADPIKATIQAPQARSAQCLSHPW-----LKSMPABEAHPINTKQLKELLAR 441
 6846 NFSYDAGDVFVKLLTEIPVSRUTVDEALDHPWINDKUKTEP-----LSADTLREFKYQ 6899
 442 SRQSRSLMSYKILVMRS-----IPELLRGP-----PDSPLGVARHLR----- 481
 6900 HKWLE-----RVFYQQTPEQILEAILGATAQAQONAPVAPEGRRPAEIVDYLRIOPK 6954
 482 -----DTGGSSSSSSSDNE-----LAPPARAKSLPSP-----V 511
 6955 KPPPTVEYVQPRKEHPPFIDFGQLIDGADPDREPGTGFEGPHRQPPQIPPPQORPQA 7014
 512 THS-----PLTHPRGFLRPSASLPPEAEASERSTEAPAPPA 547
 7015 ANDSRHEQOPQHQGPQRIPVQYGRPLVDR-YLNDPSHRPSSLDADAPPYVDYKGNFV 7073
 548 SPEGAGPAAQGCVPVPHSVIRSLFYHQAGESPEHAGALAPGSRPHRPHRLHKKGYIAGA 607
 7074 HFDKYGRPMA-----PONLEKRKLIPQDKGETPSHS--KKEKTQHPVATPIL----- 7119
 608 LQG-----LREPLMHHRVLEBAAR-----EQATLLAKAPSPETALRLPASGTH 652
 7120 SPGGDQQQKQIPMRMTRGERRIEEIANRILSDISEGSIAGSLASLED-FEIP----- 7173
 653 LAPGHSLSLHDSPTSPFPSSACGAQRLPSAPSGGADIRMGHPQGSKQLPSTGGHPG 712
 7174 ----KDFQVEASEPSTTLTPE-----VTIRE----- 7196
 713 TAQPERPSDPSFMGQAPAPCHPKQGSAPQEGCSPPHAPVACPFGSPFGSCKEAPLVPS 772
 7197 TIPKTFPSPTSPQKSVP----- 7214

QY 773 PFLGOAPPAPAKAS-----PPLDSKM--GPGDISLP-GRP--KPGPCSPGSA 817
 DB 7215 ----QPGGLLIIPAKYISDSIILAGIIPADKKVLEDAENDPSPVPGAPLPLEGLHSGDLTI 7270
 QY 818 SOASSQVSSSLRVGSSQVTEPGPSLDAEGWTQEAEDLSDSSTTLQRPQEQVMTMRKFSLG 877
 DB 7271 DTTASGL--IKVTSAINLSPNK-----SPRSTPGTKSPVLSPRQHSMEVLIAT 7322
 QY 878 GRGGYAGVAGYCTFAFGCGDAGMLQGGPMWARIANAVASQSEEEEOBEAEASQSEBQQA 937
 DB 7323 KRQ-----KPGFLPFGEL-----AEDIDDEAFMDRRKQ---- 7352
 QY 938 RAESPLOVARSARPEVGRAPTRSSPEPTWEDIGQVSLVQIRLDSGDAEAAAD----- 990
 DB 7353 -----VKPKDHGDENDFKDEKLERLEK 7373
 QY 991 -----TISL-DISEVDPA--YLNLSDL-----YDIKVLPEE--PMI-----FRKVPKSAQPE 1032
 DB 7374 DKNRRTVNLDLDDKVRPSAFYKDDSDFGHPGXDIIDATPMDSHYQIGDPDITLMAARGAEN 7433
 QY 1033 PPSMAEELABELFPEPTWPPWPGELG-PHAGLBITSEEDVDALLABAAVGRKRWKSSPER 1091
 DB 7434 SRVNVREELFGNGAPTvk-QGFLGVNRNDITVRERRYTD-ILRETTQGLEPKSHEQST 7491
 QY 1092 SLPHFPRHLPLDEPAELGLRVRVKAIVSHIRILKGRPEGLEKEGPPPKKPGCLASFRLS 1151
 DB 7492 AL-----LQKAPSATAIERIKADIEKVTPC-----ATKNDGDTGF--- 7526
 QY 1152 GLKSNDRAPTELRELSDETWLQSVTLACQVSAQAPAAQATWSKQCAPLESSESRVLISAT 1211
 DB 7527 -----AIFTRALRDVLRKNQPAIFECANVASAPAKVTDWDFQKILESNDRVITEQ 7579
 QY 1212 LKNFOLLTILVVAEDLGVYTCVSNALGTVTGTVLRKAERPSSPCPDIGEVAD-GV 1270
 DB 7580 -NNVARLIINHAAPYDLGEVCTAINNEYGTDKSSCLISGETPSRPRPE-AELSSDTEI 7637
 QY 1271 LLVWKPVESVGP-----VTYIVQCSL-----EGSWTTLASDIFDCCYLTSLSGRGTYT 1320
 DB 7638 FIOWEAPE--GPTYLEGITYRLEVRVAGNDHGDHPITVSEKIDDESIVVKHLSPLGIYQ 7695
 QY 1321 FRACVSKAGMGYS-----SPSEQV-----LIGGPH 1348
 DB 7696 FRVTAQNGFGLPLSLSSRIVQTHGKAPKLQIDVLKSBIRLNVSMPOKSTNQLGISE 7755
 QY 1349 LASEESQGBSA-----QPLSTKTFAQTQIRGRFSVVRCWEKAS-GR-LA 1396
 DB 7756 -ESEDSEARTANEDMKNLQJQTDPTGRFQIGGLKFKGRFSVIRDAVDSITTEGHAHCA 7814
 QY 1397 AKIIPYHPKDKTAVLREYALKGLRHPHQAALHAAYLSRHLVLIILELSCGPELLCLAE 1456
 DB 7815 VKI--RHPSE--AISEYESLRDQGHENQVRLIAAFNNNSFLYLLSERLY-EDVFSRFVF 7869
 QY 1457 RASYSSEVQDYVQWMLSATQYLNQHILHLDRSENMIITEYN--LLKVVLDLGNASQLS 1514
 DB 7870 NDYYTEEQVALTWRQVTSALHFLFKGIAHLVDVNPINIMFQSKRQSVWVVLVDGCRQKVS 7929
 QY 1515 QEKVLPDKFKDYLET--MAPEL-LEGQGAVPQTDIWAIGVTAIFMLSAEYPVSSSEGARD 1571
 DB 7930 -GAVKPYD-----FDIKWASPEFHIPETPVTVQSDMWGMGVVTFCLLAGFHPFTSE--YD 7981
 QY 1572 LQRLRLKGLVRLSRCVAGL-----SGGAVAFLESTLCAQPMGRCPCASSCLQCQWLTTEGP 1626
 DB 7982 REEBIKENVIN-KCDNPLIPVNASOECSFATWALKKSPVRNRWTDALSHKFLSDPS 8040
 QY 1627 ACSRPAPVTFPTARLR 1642
 DB 8041 MVRRESIKYSASRLR 8056
 RESULT 7
 Q8BZF4
 ID Q8BZF4
 AC Q8BZF4;
 PRELIMINARY; PRT; 799 AA.

DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical tyrosine protein kinase/serine/threonine protein
DE kinase/eukaryotic protein kinase/fibronectin type III domain
DE containing protein (Fragment).
GN AREG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [(1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RC MEDLINE=22354683; PubMed=12466851;
RA the FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RL ENBL; AK035543; BAC29098.1; -.
PIR; PT0566; PT0566.
DR MGD; MGI:109282; Areg1.
DR GO; GO:0005824; F1ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG-C2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 799 AA; 86397 MW; A15C45E37FB1A42 CRC64;
Query Match 9.4%; Score 818; DB 11; Length 799;
Best Local Similarity 29.4%; Pred. No. 5e-44;
Matches 205; Conservative 83; Mismatches 234; Indels 176; Gaps 7;
QY 1140 RKXPGLASFRLSGL-----KSWDRAFTFLRELSDETIVLQSVTLACQVSAQPA 1188
DB 93 RKDKGLSQNLSSVQEDLGHQVPSFSDFPVFHKLKQVQLGEAATLLCLPAACPA 152
QY 1189 ACATWSKDGAPLESSRVLISATLKNFQLTLVWVAEDLGVYTCVSNALGTVTTGVL 1248
DB 153 PRISWMKQSLRSESVVIVSKDGRQLSLPRAGKRHAGLYECSATNVLSITSCVT 212
QY 1249 RKAERSSPCPDIGEVADGVLLVWKPVESGPTVYVQCSLEGG-NTTASDIEDCC 1307
DB 213 AVARIFGKLAPEVPQTHDTALVWVKPGDGRAPCTYTLERRVDGESVWHPVSSGIPDCY 272
QY 1308 YLTLSLRGTTVTRTACVSKAGMGFPYSPSPQVLLGG----- 1345
DB 273 YNVTQLPVGVTVFRVACNCRAGQGFPSSEKVFIRGTPDSPAAPAAAPRDPVTSPT 332
QY 1346 -----PSHLA----- 1350

DB 333 RAPPDPSPTSLAPTALAPPASQASTLSFSTSSMSANQALSSLKAVGPPPTPPKRRGL 392
QY 1351 -----SEERS-----QG----- 1357
DB 393 LATQAEPSPPSIVVTPSEPRSPVDPDCTGLTPTSPQGVKAPSPSTSLYMTSVFSAPPA 452
QY 1358 -----RSAQPL-----PSTKTFAPQTQIQGRFSVV 1383
DB 453 PQAPAPEPPEPTKVTVRSLSPAKEVSSPTBTTLRQGPQPKPTFLEEKARGFGV 512
QY 1384 ROCWEKASGRALAAKIIIPHPKDKTAVLRVEALKRHLRHPHQAQHAAYLSPRHVLVILE 1443
DB 513 RSCRENATGRTEFAKIVPAAGKRVLOEVEVTLTLLHRELMSLHEAYITPRVLVIAE 572
QY 1444 LCGPELLPCLAEPRASYSESEKVKYLMQMLSATYLNQHLHLDLSENMITEYNLLK 1503
DB 573 SCGNRELLCGLSDFRYSDDVATVVQLQGLDYLGHVHLHDIKPDNLLAADNALK 632
QY 1504 VVDLGNAGLSOEKVLPSDKFVDLETMAPELLEGQGAVPOTDIWAGVTAFLMSAEP 1563
DB 633 IVDFGSAQPNPQALKPLGHTGTLEFVAPENWVGDPGSGATDINGAGVLTYYIMSGVSP 692
QY 1564 VSSEGARDLORGLRGLVRLSRCYAGLSGGVAFLRSTLCAPQWGRPCASSCLOCPLWTE 1623
DB 693 FYEPDPPQTEARIVGGREFDAFQLYPNTSQSATLFLRKVLVHPWSPRSLOCLAPHLQD 752
QY 1624 EGPACSRPAPVTFPTARLVRVFRNREKERRALLYKRHL 1661
DB 753 AYLMKLRQTLTFTTNRLKEFLGEGRRRAEAATRHKV 790
RESULT 8
Q7Z119 PRELIMINARY; PRT; 1393 AA.
AC Q7Z119;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Uncoordinated protein 89, isoform C.
GN UNC-89.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodexinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [(1)_TaxID=6239;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [(2)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "The sequence of C. elegans cosmid c09D1.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [(3)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [(4)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003131; AAP68959.1; -.
SQ SEQUENCE 1393 AA; 156296 MW; F947C1BCE4EB9CE9 CRC64;
Query Match 8.0%; Score 701.5; DB 5; Length 1393;
Best Local Similarity 22.3%; Pred. No. 3.7e-36;

Matches	364;	Conservative	212;	Mismatches	580;	Indels	475;	Gaps	691;
QY	211	DILAASHPLVTGLLQDPETRKTLILILELCSSEILLDRYKGVVTEAEVKVYIQOLVE	270						
Db	14	DGLSSLAHPGV-----EIAE---PKGVNRETQVRFVRLQLL	47						
QY	271	GLHYLHSHGVHLHDIKPSNLMVHPAREDIKICDFGFAQNIIPAEQFQSOYQSPFVSPE	330						
Db	48	ALKHMDLRIAHLDLPEFILL---ODDKLKLADFQARLLRGLITGSIKSGSPFVSPE	104						
QY	331	IIQONPVSEASDIWAMGVISYLSLTCSPPAGESDRAATLLNVLEGRVSWSSWAAHLSDE	390						
Db	105	IVRSYPLTLATDMWSTGVLTVYLLTGLSPFHGNDNETLANVDSQCFD--SSPL-GNFSYD	162						
QY	391	AKDFIKATLORAPQARPSAQCILSHWF---LKSMPAEAEHPTNKQKFLARLRQWR	446						
Db	163	AGDFVKLLTEIVSRLTVDDEALDHPWINDKLUKTEP-----LSADTLREFYIQHKWLE	216						
QY	447	SLMSYKSILVMRS--IPELLRGP-----PDSPSLGVARHLCR-----	481						
Db	217	-----RRVFQOFTPSQILLEAILGPATAQAQNAVPAPGERRPAEIVDYLRIQKPPPT	271						
QY	482	-----DTGSSSSSSSSDNE-----LAPFARAKSLPPSP-----VTHS---	514						
Db	272	VEYVQPRKEHPFIDFEGQLIDGADFDPREGTGFEGHRQPQIIPQOPRNQAAHDSR	331						
QY	515	-----FLLHPRGFLRPSASLPEEAESAERSTAPAPAPASPEGA	552						
Db	332	RHEQQQHQCCQPQIPVDVQVGRPLVDPR-YLNDP-SHRPSSLDDAPFVYKYGNPVHFDKY	390						
QY	553	GPAAQCGVPRHVSIRSLFYHQAGESPEHGALAPGSRHRPARRRHLLKGYVIGALPG--	610						
Db	391	GRPWA---PQNLEKRLIPQDKGETPSHS--KKENTQHPVATPIL-----ASFGGD	436						
QY	611	-----LRPLMEHRVLEBEAR-----EEQATLLAKAPSETALRLPASCOTHLAPGH	657						
Db	437	QQQKIPMRMIRGERIEIEEIANRILSDISEGSIAGSLASLED-FTFP-----K	486						
QY	658	SHSLEHDSPTPRPSSACQALPSAPSGGAPIRDMGHPOGSKOLPTGSGHPTAQPE	717						
Db	487	DFQVEASESTPILTEP-----VTIRE-----TIKPK	513						
QY	718	RPSDPSPWGQAPAFCHPKQGSAPQEGCSHPHAPVAPCPGSPFGSCKEAPLVSSPFLGQ	777						
Db	514	TPSPTSPQKSFPV-----PPLDSKM---GPGDISLP--GRP--KPGPCSSPGSASQASS	822						
QY	778	PQAPPAPAKAS-----PPLDSKM---GPGDISLP--GRP--KPGPCSSPGSASQASS	822						
Db	528	PQGLLIPAKVYSDSILLAGLPADKVKVLEDAENDPSIPVGAFLFEGHSGSDITITTSA	587						
QY	823	SOVSSLRVSGSVQGTTEPGSLDAEGMTQEAEDLSDSTPTLQRPQEQVTMRKFSLGRGGY	882						
Db	588	SGL--IKVTSPLAINLSPNK-----SPRSTPTGKSPVLSPRQEHSMSEVLATKRG--	637						
QY	883	AGVAGYGTFAFGDAGGMLQGGPMWARIANAVSQSEEEQEEARASQSEEQEAREASP	942						
Db	638	-----KPGFLPPGEL-----AEDIDEDAFMDRKKQ-----	664						
QY	943	LPQVSARPVVEVGRAPTRSSPEPTPWEDIGQVSLVQIRDLSDGAEAAD-----	990						
Db	665	-----VKPKDHGDGENDFKDKERLEKDKQVR	690						
QY	991	TISL-DISEVDPA--YLNLSDL-----YDIKYLPE--FMI-----FRKVPKSAQPEPPSPM	1037						
Db	691	TVNLDLDRKPSAFYKDDSDFGHPGYDIDATPMWDSHYQIPDPTYLMAARGAFAFNRSVRN	750						
QY	1038	AEEELAFPEPTWPWPGLG--PHAGLITITSESDVDALLAAVGRKKRWSSPSRSLFHF	1096						
Db	751	YREELFGMAFTVK--QGFLGVNRDITVRRERYTD--ILRETTQGLEPKSHEQSTAL---	805						
QY	1097	PGRHLPLDEPAELGLRRVKASVEHILGRPEGLEKSGPPPKPGLASFRLSGLKSW	1156						
Db	806	-----LOKAPSAVAIERIKADIEKVTPC-----ATPKNDGDTF-----	838						

RESULT 9

AC	Q9W1D6	PRELIMINARY;	PRT;	1226 AA.
AC	Q9W1D6;			
DT	01-MAY-2000	(T=EMBLrel. 13, Created)		
DT	01-OCT-2002	(T=EMBLrel. 22, Last sequence update)		
DT	01-OCT-2003	(T=EMBLrel. 25, Last annotation update)		
DE	CG30171	protein (GH20492p).		
DE	CBEST:HL01080	OR CG18019	OR CG18020	OR CG18021
OS	Drosophila	melanogaster (Fruit fly).		
OS	Eukaryota;	Metazoa;	Arthropoda;	Insecta; Pterygota;
OC	Neoptera;	Endopterygota;	Diptera;	Brachycera; Muscomorpha;
OC	Ephydroidea;	Drosophilidae;	Drosophila.	
OC	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkely;			
RC	MEDLINE=20196006;	PubMed=10731132;		
RA	Adams M.D.,	Celniker S.E.,	Holt R.A.,	Evans C.A.,
RA	Ananitskies P.G.,	Scherer S.E.,	Li P.W.,	Hoskins R.A.,
RA	George R.A.,	Lewis S.B.,	Richards S.,	Ashburner M.,
RA	Sutton G.G.,	Wortman J.R.,	Yandell M.D.,	Zhang Q.,
RA	Wan R.C.,	Rogers V.-H.C.,	Blazej R.G.,	Champe M.,
RA	Brand K.H.,	Doyle C.,	Baxter E.G.,	Helt G.,
RA	Abiril J.F.,	Agbayani A.,	An H.-J.,	Andrews-Frannkoch C.,
RA	Bailew R.M.,	Basu A.,	Baxendale J.,	Bayraktaroglu L.,
RA	Beeson K.Y.,	Banos P.V.,	Berman B.P.,	Bhandari D.,
RA	Borkova D.,	Borchan M.R.,	Bouck J.,	Brokstein P.,
RA	Burtis K.C.,	Busam D.A.,	Butler H.,	Cadiou E.,
RA	Cherry J.M.,	Cawley S.,	Dahlke C.,	Davenport L.B.,
RA	de Pablos B.,	Delcher A.,	Deng Z.,	Mays A.D.,
RA			Dew I.,	Dietz S.M.,

QY 706 STGGHP---GTAQPE---RPSDPSWPGQAPFCHPKQGSAPQGCSPHPAVAPCPGSG 758
Db SKYLHPDYELGLIQSESHYQYGDPT : : : : : 557
QY 759 PPGCKEAPLVPSPFFLQGPQAPPAPAKASPLDSKMGPCDLSLPORPKPGSCSFGSGAS 818
Db : : : : : 573
QY 819 QASSSQVSSLRVGSSQVGTBPQPSL---DAEGWTOEAEDLSSTPLORPQSQVMTWRKFS 875
Db : : : : : 574
QY 876 LGGRGYAGVAGYGTAFAGDAGGMLGQGPWARIAMVAVSQSEEBEQAESQESHQ 935
Db : : : : : 606
QY 936 EAAESPLQVQSPVPEVGRAPTRSSPEPTMEDICQVSLVQIRDLSDGAAADTISLD 995
Db : : : : : 614
QY 996 ISEVDPAVNLSDYDKYLPFFEMFIRKVPKSAQPEPPSPMAEELABFPPTWPGCE 1055
Db : : : : : 627
QY 1056 LGPHAGLEITEESDVEDVALLAAEAVGRKXKWSRSLFFHFGHLPDPAELGLRERV 1115
Db : : : : : 639
QY 1116 KASVEHSIRILKRGPEGLEKGPGRKPKGLASPLSLKSNDRAPTFRLSDETVVL--- 1173
Db : : : : : 655
QY 1174 QSVTLACQVSAQAQATWSKDGAPLESRRVLISATLKNFQLTILVVVAEDLGVYTC 1233
Db : : : : : 676
QY 1234 SVSNALGTVTTGVLKAEAPSPSPDCEVYVADGVLLWKPVSYSQPV---YIVQCS 1290
Db : : : : : 736
QY 1291 LEG-GSWTILASIDFCCLYTLKSLRGVTVTRTACVSKAGMPYSSPSEQVLLGSPS-- 1347
Db : : : : : 796
QY 1348 -----HLASBESQGRSAQP-----LPSTKT-----FAFQTOIQRG 1378
Db : : : : : 856
QY 1379 RFSVROCWEKASGRALAAKIIYPHKDKTAVLRVEALKGLRHPLAQLHAAY--LSPR 1436
Db : : : : : 915
QY 1437 HLVLILELCSGPPELLCLABRASYSESVKQYLMQWLSATQVYLHNOHILHLDLRENMII 1496
Db : : : : : 975
QY 1497 TEYNLL--KVVDLGNQSLSQ--EKVLPSPDKPKDYLETWAPELLBGQAVPQTDIWAIGV 1552
Db : : : : : 1035
QY 1553 TAFMLSAEYVPVSESGARDIQRLKGLVRLGRVAGLGGAVAFRLSTLCAQWGRPCA 1612
Db : : : : : 1091
QY 1613 SSCIQCPWLTEEGPACSRPAPVTFPTARLRFV 1644
Db : : : : : 1151

RESULT 10
Q9W1D5
ID Q9W1D5
AC Q9W1D5;

PRELIMINARY; PRT; 3197 AA.

01-MAY-2000 (T-EMBLrel. 13, Created)
01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
CG30171 protein.
GN BEST:HL01080 OR CG3901 OR CG18019 OR CG18020 OR CG18021 OR CG30171.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Echioidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Berkely;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.R., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.P., Agbayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,
RA Balles R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherty J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Bunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeiler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.N., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RP Celiniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferrera S., Fouse E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Parasag V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,

RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome";
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN
 RN SEQUENCE FROM N.A.
 RP
 RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN
 RN SEQUENCE FROM N.A.
 RP
 RP FlyBase;
 RA
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AE0033462; AAF47134.2;
 RL FlyBase; FBgn0040499; BEST:HL01080.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPRO03961; FN_III.
 DR InterPro; IPRO03599; IG_III.
 DR InterPro; IPRO07110; IG-like.
 DR InterPro; IPRO03598; IG_c2.
 DR InterPro; IPRO00719; Prot_kinase.
 DR InterPro; IPRO02290; Ser_Thr_kinase.
 DR InterPro; IPRO01245; Tyr_kinase.
 DR Pfam; PF00041; fn3_2.
 DR Pfam; PF00047; ig_15.
 DR Pfam; PF00069; Pkinase; 2.
 DR Prodom; PD0000001; Prot_kinase; 2.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00409; IG; 21.
 DR SMART; SM00408; IGC2; 18.
 DR SMART; SM00220; S_TKc; 2.
 DR SMART; SM00219; Ty-Kc; 1.
 DR PROSITE; PSS0835; IG-LIKE; 14.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 2.
 DR ATP-binding; Immunoglobulin domain; Transferase.
 SW SEQUENCE 3197 AA; 361286 MW; F4E74689B45A936 CRC64;

Query Match	8.0%	Score	695.5;	DB	5;	Length	3197;
Best Local Similarity	18.5%;	Prod.	No. 3.1e+35;				
Matches	305;	Conservative	209;	Mismatches	513;	Indels	625;
						Gaps	41;
QY	63	VOATGGTAPEAII	EGDPPSVTWYKDSVQLVDSTRLS--QQOEGTYSVLVRHVASKDA	121			
DB							
QY	2057	IMARKNG--RIEAKL	VGIPUPEVHVPKDWKPIVDSRSIKISSYDDPIVLSIHDSIIKDG	2114			
DB							
QY	122	GVYTCLAQNTGGQVL	CKAELLVJGGNEDPDSKQSHRR-----KLHFSVFKVEBIG	172			
DB							
QY	2115	GLYSISARNTAGSI--	STSVTHIEENERDQYIKTYGRHPYVRSKQLRYQDKYDIGBELG	2172			
DB							
QY	173	RGVFGFVKRVQHKNTL	CAAKEIPRSTRCAQAVRERDILAAISHPLVTGLLQDFQFTRK	232			
DB							
QY	2173	RGTOGITVTHAVERS	SGDNAAKIMYGRPELRPFMLNELEMNMTFNHKLIRPYDAYDTR	2232			
DB							
QY	233	TLILILELCSSEELL--	DRLYRKGVVTEAEVKYVQQLVEGLYHLHSHGVYHLDIKPNIL	291			
DB							
QY	2233	SVTLIMELAAAGGEL	VLRLNLRDYYTERDIAHYIQTLMGLEHMEVGMVGHMTIKDLL	2292			
DB							
QY	292	MVHPAREDIKIDFG	FAQNTTPAELQFQSGPSPFVSPBEITQONPVSASDITWAMGISY	351			
DB							
QY	2293	ISVVGDDIIKVSDF	GLSRKINRNLSTLDYGMPEFVSVVYNKEGVNPSHDMWTVGLTY	2352			
DB							
QY	352	LSLTCSPFAGEDRAT	LLNLVLEGRVSWSPAAHLSEDAKDFIKATLQAPQAPPAQAQ	411			
DB							
QY	2353	VLLGGHNPFGIDRE	TLTKIREGRWDFKDEIWTHTISDDGRDFTSRLLSYSPERMDVKT	2412			
DB							
QY	412	CLSHPWFLKMPAAE	AHFINTKQKELLARSQWRLSYKSIILVWRSIPPELLRPPDPS	471			
DB							

QY 1497 TEYNLL--KVDLGNALQSLQ--EKVLPDSKFKDYLETMAPELLEGOGAVPQTDINAIGV 1552
 Db 3006 ASVRSIQVLVDGSGAKVKNLGMKVTGCGS---LDFQPEMINDEPIFPQSDINSLGA 3061
 QY 1553 TAFIMLSAEYVSVSEGRADLQGRKGLVLRSCYAGLSGAVAFURSTICAPQWGRPCA 1612
 Db 3062 IYLLLSGCSFPGADEYETKQTSFVRYRPFENLFKEVTPFATRFIMLLFKRHPTRKPYT 3121
 QY 1613 SSSLCQPLWTBEGFACSPAPVTPPTARLRFV 1644
 Db 3122 EDCLEHRLWSSDYVVRKERRAIFLGSRLKTF 3153

RESULT 11

ID Q80W23 PRELIMINARY; PRT; 1042 AA.
 AC Q80W23;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Trio protein (fragment).
 GN TRIO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Olfactory epithelium;
 RX MEDLINE=22388257; PubMed=2477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grunwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerk A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006488; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR000219; RhoGEF.
 DR InterPro; IPR002290; Ser thr pkinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00169; PH; 1.

Pfam: PF00069; pkinase; 1.
 DR ProDom; PD000003; Prot_kinase; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS001010; DH 2; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00003; PH DOMAIN; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_SI; 1.
 FT NON_TER 1
 SQ SEQUENCE 1042 AA; 113513 MW; 1CA464B9AD694032 CRC64;

Query Match 6.6%; Score 580.5; DB 11; Length 1042;
 Best Local Similarity 33.7%; Pred. No. 1.6e-28;
 Matches 141; Conservative 75; Mismatches 192; Indels 11; Gaps 6;

QY 53 PPSMQVTIEDVQAQTGTAQFEALIEGDPQPSVTW---YKDSVQLVDSTLSQQQGGTTY 109
 Db 627 PPEVIFLSEVTGTTGTVVRCVGRPKASITWGPBHTLNDDHYSISYSDIGEA- 685
 QY 110 SLVLRHVASKDAGVYTCLAQTGGQVLCKAELLVLGDNPEPDSEKSHRRKLHSFYEVKE 169
 Db 686 TLKIIGVSTEDDGIYTCIAVNDMGSSASSASLRVLGFGS--DGI VVTWKDNDFAYSEVA 743
 QY 170 EIGRGVFGFKRVQHKNLCKAETPLRSRTEAQAAYRERDILAAALSHPLVTGLDQSE 229
 Db 744 ELGRGRFAVVKCKDQKQKAVATKFNKMKMGKQDVTHLGLQLQLPPLVSLDQTE 803
 QY 230 TRKTLILILCSSEELLRLYRKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKPSN 289
 Db 804 TPTSIVLVLEMAQGRLLDCCVVRGSLTEGKVRHLGEVLEAVRYLHNCRIAHLDLKPEN 863
 QY 290 ILMVHP-AREDIKICDGFQAQNTPAELQFSQVGSSEFVSPELIQONPYSEASDITWAGV 348
 Db 864 ILVDQSLAKETIKLADFGDAVQNTTYYIHLLGNPEFAPELIILGNVSLTADTWSVG 923
 QY 349 ISYLSLTCSPFFAGESDRATLNLVLEGRVSWSPMAHLSEDAKDFIKATLQAPQAPES 408
 Db 924 LTVYLLSGVSFFLDDSVETELNLCRLDSFPEDYFGVSKAKSEVCFLLQEDPAKPRS 983
 QY 409 AAQCLSHPWPF-LKSMFAEAEHAFINTQLKFLARSWQ---RSLMSYKSLVWRSIPEL 463
 Db 984 AALALQEWLQAGNGSGKGTGVLDTSLRTLSFTIERRKHONDVPRISIKNFLOSRLPRV 1042

RESULT 12

ID Q961U1 PRELIMINARY; PRT; 1721 AA.
 AC Q961U1;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE GH07636P.
 GN BT OR CG1479 OR CG10285 OR CG32019.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Prounanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.W., Celisner S.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDJ databases.
 DR EMBL; AY047563; AAK77295.1; -.
 DR FlyBase; FBgn0005666; bt.

```

DR GO: 0005524; F-ATP binding; IEA.
DR GO: 0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: 00016740; F:transferase activity; IEA.
DR GO: 0000469; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR003962; FnuII subd.
DR InterPro: IPR003961; FN III.
DR InterPro: IPR008957; FN III-like.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig c2.
DR InterPro: IPR006162; Ppantne S.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00041; fn3; 3.
DR Pfam: PF00047; ig; 8.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEII.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00408; IGC2; 5.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS50835; IG LIKE; 8.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Immunoglobulin domain; Repeat; Transferase.
KW SEQUENCE 1721 AA; 193936 MW; 500774B1E81EFB93 CRC64;

Query Match 6.3%; Score 552; DB 5; Length 1721;
Best Local Similarity 16.6%; Pred. No. 2,3e-26;
Matches 313; Conservative 172; Mismatches 453; Indels 748; Gaps 51;

QY 48 PALPGPPSMQVTTEDVQAQTGGTAIEGDPQPSVTVYKDSVLVDSTRLSQQQSGT 107
DB 430 PQAASPLIVKPLRDANCIONHNAQTCTTINGVPKPTISWYKGAREISNGARYHMYSEG 489

QY 108 TYSILVLRHVASKDAGVYVTCIAQTNGOVLCCKALLV----- 143
DB 490 NHFLINDVFGEDEAVCYRANVKAQAKSTRATLAIATAPKLVNPFPRDTAYFDKGENV 549

QY 144 ----- 143
DB 550 VIKIPFTGPKPRIHWVRDGENIESGGHYTVVEKRVHAVLIIRDGSHLSDGPPRYTAENE 609

QY 144 LGGD-----NEPDS-----EKQSH--- 157
DB 610 LGSDTAIQVQISDRDPFRPPLIESIGTSLSKWAPVWDGSDITNYYVERREHPLS 659

QY 158 ----- 157
DB 670 SWIRVGNTRTSMVAVSLTPGEYDFRIADNVVGRSDASDTSLTIKTESVKKKPIERK 729

QY 158 -----RKLHSPYEVKEISGVEFG 177
DB 730 WEIDANGKRLRGKADGPKVDYSYVFDIYSKFEVQPVIEISQQSYVDYRDILEISGTGAFG 789

QY 178 FVKEVQHKNGKILCAAKFIPL-RSRTRAQYRRERDILAALSHPLVTGLLDQFTRKTLIL 236
DB 790 VVHCRERSTGNIIPAAKFI PVSHSVKDLIRREIDINQLHHQKLIHLHDAFEDDDDEML 849

QY 237 ILELCSSEELLDRLYRKG-VVTEAEVKVYIQQLVEGLHYLHSHGVHLHDIKPSNILMVHP 295
DB 850 ILEFLSGELPERITAEGYVTEAEVNYRQICEGRHMHBEQNIHLHDIKPENIMCQTR 909

QY 296 AREDIKICDFGAQNIPTAELOFQSQVGPSPFVSPETIQQNPVSEASDIWANGVISYLSLT 355
DB 910 SSTNVKILDFGLATRLDPNEVVKITGTGAAPAEIVNRFVPGYTDWATGVLSYVLLS 969

QY 356 CSSPFAGESDRATLLNLEGRVSWSSPVAHLSEDAKDFKATLQAPAPSPSAQCLSH 415
DB 970 GLSPFAGNDVQTLKVNKACDWDPDFVESFKYISEAKDFIKLLVRNKRKMTAHECLLH 1029

```

```

DB 1701 AVSGKA 1706
|||||
AC Q9V4F7 PRELIMINARY; PRT; 8943 AA.
ID Q9V4F7
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CG32013-PA.
GN BT OR CG32019.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis J.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S.A., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Houston D., Houston K.A., Howland T.J., Wei M.H., Ibegwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2183-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Mirza S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celnikier S.E.,
RA Clamp M.E., Drysdale R.A., Emert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kronmiller B., Marshall B., Milburn G.H., Richter J.,
RA Ruso S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome."
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003843; AAEP59316.4; -.
DR FlyBase; EGN0005666; bc.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000577; FGGY kin.
DR InterPro; IPR003962; FNIIT subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG c2.
DR InterPro; IPR006162; Pplantne S.
DR InterPro; IPR000719; prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR008271; Ser Thr_kin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00041; fn3; 39.
DR Pfam; PF00047; ig; 26.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00060; FN3; 40.
DR SMART; SM00409; IG; 39.
DR SMART; SM00408; IGc2; 29.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00933; FGGY KINASES 1; 1.
DR PROSITE; PS00835; IG LIKE; 30.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_S1; 1.
SQ SEQUENCE 8943 AA; 1001276 MW; 786EA35A4E62ED47 CRC64;
Query Match 6.3%; Score 552; DB 5; Length 8943;
Best Local Similarity 18.6%; Pred. No. 2.6e-25;
Matches 313; Conservative 172; Mismatches 453; Indels 748; Gaps 51;
QY 48 PALPGPPSMQVTTEDVQAQTGGTAQFAEALIEGDPQPSVTWYKDSVLVDSTLSQQCEGT 107
DB 7652 PQASPPPLIVKPRDANCIONHNAQFTCTINGVFKETISWYKAREISNGARYHMYSEGD 7711
QY 108 TYSILVLRHVASKDAGVYVTCIAQNTGGVCLCKAELIV----- 143
DB 7712 NHFLINDVFGEDADEVYCVKAVNKAGAKSTRATLAINTAPKLVPPRFRDTAYFKGENV 7771
QY 144 ----- 143
DB 7772 VIKIPTGPKPIHVRGNIENIESGHYTVVEKHAHLIIRDSGLDGGPRYTAENE 7831
QY 144 LGGD-----NEPDS-----EKQSH----- 157
DB 7832 LGSDTAILQVSDRDPDPPLIESIGTESLSLWKAPWGGCDITNYVVEREHPIS 7891
QY 158 ----- 157
DB 7892 SWRVGNTRFETSMVSGLPFGKEIDYPRIFADNVYGRSDASTSLTKIKESVKKPIERK 7951
QY 158 -----REKLKHSFYVEKKEIQRGVFG 177
DB 7952 WEIDANGKRLGRKADGVKDYVFDIYSKVPQVPEISQQSVYDVIDILSEIGTGAPG 8011
QY 178 FVKRVQHKGNKILCAAKFTPL-RSRTRAQAYERDILALSHPLVTGLDQETKTLIL 236

```


8012	VHRCRSTGNIFAAKTFIVSHSVSKDLIRREIDIMNQLHHQKLIINHDAFEDDDDEMIL	8071
QY	237 ILELCSBELLDRLYRKQ-VVTEAEVKYIIQOLVEGLHYLHSHGVLHDIKPNSILMVHP	295
DB	8072 ILEFLSGGELFERITAEVYVWTEAEVINYRQICEGRHMEQNIHLDIKPENIMCQTR	8131
QY	296 AREDIKICDFCAQNITPAELQFSQYSGPEFVSPELLICQNPVSEASIMWANGVLSYLSLT	355
DB	8132 SSTNVKLDLDFGLATRLDNEVVKITGTGAEPAAPEIVNREPEVGYTDMWATGVLUSYLLS	8191
QY	356 CSSPFGBSDRATILNVLEGVSWSSPMAHLSEDAKDFIKATIQRAPQARPSAAQCLSH	415
DB	8192 GLSPFAGNDVQTLKNVKACDWDDEVSFKYISBEAKDFIRKLVRNKEKRWMTAHECLLH	8251
QY	416 PWFLKMPABEBAHINTKQLKFLARSRWQHSLSYKSIIV-----MESIPELLRGPP	468
DB	8252 PWLTGDRSAMQE-INRD--RYLAYREKLRRKYEDFERFLLPICRLSEYSLKXLL-----	8304
QY	469 DSPSLGVARHUCRDTGGSSSSSDNELAPFAKSLPSPFVTHSPLLHPRGFLRPSAS	528
DB	8305 -----MEKYIHD-----AVFDRQAA-----PRFVIRPS-----	8329
QY	529 LPEZAEASERSTEAPAPASPEGAGPAAQCQVPRHSVIRSLFYHQAGESPEHGALPGS	598
DB	8330 -----SQCFCYGSQSV-----KFCYRC-----	8345
QY	589 RHPARRHLKGYIAGALPCLREPLMEHRVLEBEEAREEQATLLAKASFE-----	641
DB	8346 -----IAIATPTL--TWSHNNIE-----LRQYKFKRYVGDG	8376
QY	642 -----TALRLPASGTHLAPGHSHSLEHDS-----PSTPRPSEACEAQRULPSAPSGGAPI	692
DB	8377 YVFIINRVKLDORGEYIIRAENHYGSRREVVLNVQLPKEQ-----PRYRTESTPV	8428
QY	693 RDMGHPQSGKOLPTGGHPGTAQERPSPSQWCPQAPFCHPKQGSAPQGGCSFHAVAP	752
DB	8429 R-----RREPLPYTFW-----QB-----	8441
QY	753 CPFGSPFGSCKEAPLVSPSPFLGQOPAPAPAKASPLDSSKMGPGD-----ISLPGRPK	807
DB	8442 -----ESETAPSFTELLRPRV-----MQARDCTKLCLCLSGKPV	8475
QY	808 PGPCSSPGSASQASSQVSLRVGSSQVGTGTPGSPSLDAEQWTQAEPLDSTPTLQRPQE	867
DB	8476 P-----NVR-----WYKDRELSKYEVAMTHSDG	8499
QY	868 QVTWRKPSLGGRGYAGVAGYGTTFAGDAGGMLGQGPMMARIATAWVSQSEEBEQEARA	927
DB	8500 VVTME-----	8504
QY	928 ESQSEEQCEARAESPLPOVSARPYVCGRAPTRSSPEPTWEDIGOVSLQVIRDLSGDAE	987
DB	8505 -----IIDCKP-----SUGKYS-CRATNCHGTDE	8528
QY	988 AADTISLSDISEVDPAYNLSD--LY--DIKYLPEPFMIFRKVPKSAQPEPPSPMAEBELA	1043
DB	8529 TDCVIVIEGEVWVTEQQAHLNFYISGDRKYI-----EQIPKAPLPIVTSR	8575
QY	1044 BFPEPTWPCGELPHAGLEITEESDVALLA---EAAVGRKKWSSPSSRSLFHPFGRH	1100
DB	8576 QYTS-----SSVQNTSPQDKNVVNSNSGGINKKKYASNS---LQAPG--	8618
QY	1101 LPLEBPALGLURVRKASVEHISRLIKRPEGLEKEGPPRKKPGCLSAFRLSGLUKWMDRAP	1160
DB	8619 -----SPSRSRSATKELILFPDDSLMCKP	8642
QY	1161 TFLRELSDETIVLGQSVTLACQVSAQPAQAATWSKDGAPLESSSRVLISATLKN-FOLLT	1219
DB	8643 EFTKPLDLTLIHDEQILITCYVKGDPEPQLISNKNKGSLSSSD--ILDLRYKNXGATLTL	8700
QY	1220 ILVVVAEDLGVYTCVSNALGTVTTTGVL-----RKARPPSSP-----	1258
DB	8701 INEVFPDEGVITCTATNSGAVETCKKLTIQPLDKINRKNVAGNAPKIVSHLESRF	8760

QY	1259	-----CPDIGEYADGVLLVW-----KPVESGPTVYIIVQCSLEGSWTTLAS 1301
DB	8761	VRDGDVAVNLACRIIAGAHFD---VVHLHNNKEIKPSKDF-----QYTNEANIYRLQIA 8810
QY	1302	DIPCCVLTSLKISGGTYT-----PRTACVSKAGMPYSS-----PSEQVLL 1343
DB	8811	EIFP-----EDGGHYTCEAFNDIGESFSICTINVTVPGETKQPSFVKPFTSVSULE 8862
QY	1344	GGPHLASEBSQ-----GRSAQPIPLST-KTFAPQTQIQGRFSV-----VRQCWEX 1389
DB	8863	GGGTFPECEIDSELLNLWLKDGKPIDETPLRYSFTKDGHRYSFAVAKCNMDVDVGQYQAK 8922
QY	1390	A-SGRA 1394
DB	8923	AVSGKA 8928

RESULT 14

Q86GD6	PRELIMINARY;	PRT; 8625 AA.
AC	Q86GD6;	
DT	01-JUN-2003 (TrEMBLrel. 24, Created)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Projectin.	
GN	PRO1.	
OS	Procambarus clarkii (Red swamp crayfish).	
OC	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;	
OC	Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;	
OC	Asacidae; Cambaridae; Procambarus.	
OX	NBI_TaxID=6728;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Muscle;	
RA	Oshino T.; Shimamura J., Fukuzawa A., Maruyama K., Kimura S.;	
RT	"The entire cDNA sequences of projectin isoforms of crayfish claw	
RT	closer and flexor muscles and their localization.";	
EL	J Muscle Res Cell Motil. 0:0-0(2003).	
DR	EMBL; AB055927; BAC66140.1;	
DR	GO; GO:0005743; C:mitochondrial inner membrane; IEA.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0005488; F:binding; IEA.	
DR	GO; GO:0004713; F:protein serine/threonine kinase activity; IEA.	
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.	
DR	GO; GO:0003743; F:translation initiation factor activity; IEA.	
DR	GO; GO:0006413; P:protein amino acid phosphorylation; IEA.	
DR	GO; GO:0006413; P:translational initiation; IEA.	
DR	GO; GO:0006810; P:transport; IEA.	
DR	IncrPro; IPR003962; F:III subd.	
DR	IncrPro; IPR003961; F:III.	
DR	IncrPro; IPR008957; F:III-like.	
DR	IncrPro; IPR003599; Ig.	
DR	IncrPro; IPR007110; Ig-like.	
DR	IncrPro; IPR003598; Ig-c2.	
DR	IncrPro; IPR003596; Ig-v.	
DR	IncrPro; IPR001993; Mitoch carrier.	
DR	IncrPro; IPR000713; Prot kinase.	
DR	IncrPro; IPR002290; Ser Thr_pkinase.	
DR	IncrPro; IPR008271; Ser Thr_pkin_AS.	
DR	IncrPro; IPR001950; TIF_Suil.	
DR	IncrPro; IPR001245; Tyr_pkinase.	
DR	Pfam; PF00041; fn3; 39.	
DR	Pfam; PF00047; ig; 13.	
DR	Pfam; PF00069; pkinase; 1.	
DR	PRINTS; PR00014; FNTYPEIII.	
DR	ProDom; PD000001; Prot kinase; 1.	
DR	SMART; SM00060; FN3; 39.	
DR	SMART; SM00409; IG; 36.	
DR	SMART; SM00408; IGc2; 24.	
DR	SMART; SM00406; IGV; 3.	
DR	SMART; SM00220; S.TKC; 1.	
DR	SMART; SM00219; TykK; 1.	

DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 2.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGG2; 2.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS0835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Immunoglobulin domain; Kinase; Repeat;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 992 AA; 110185 MW; 3ABD72460983269 CRC64;

Query Match 6.2%; Score 541; DB 4; Length 992;
Best Local Similarity 24.4%; Pred. No. 5.1e-26;
Matches 140; Conservative 91; Mismatches 196; Indels 146; Gaps 7;

QY 45 SSLPALPG-----PPSMQVTIEDVQAOTGGTAPEAIIIGDPPQ 83
DB 286 SSLPVLGTSDATVKKKPAKTPPKAAMPFQIIQFPEDQKVRAGESVELFGKVTGTQPI 345

QY 84 SVTWKDSVOLVDSTRLSQQEGTYSILVRHVASKDAGVYTCLAQNTGGQVLCKAELLV 143
DB 346 TCTWKFQKQSEHMKVENSENGSKLTILAAQEHCGCYTLLENKLGSRQAQYNLTIV 405

QY 144 LG----- 145
DB 406 VDKPPDPAGTPCASDIRSSLTLSWGYSGYDGGSAVQSIEIWDSSANKTWKELATCRST 465
QY 146 -----CDNEP-----DSEKO--- 155
DB 466 SFNVQDLLPDHEYKFRVRAINVTSEPSQSELTIVGEKPEEPKDEVEVSDDEKEPEV 525
QY 156 -----SHRKLHSFVEVKEEIGRGVGFYKRYQHKGNKILCAAKFIPLSRTRAQAYR 208
DB 526 DVRTVTINTEQKVSDFYDIERLGSKGFGQVFLVEKTKRWAGKFKAYSAKEKENIR 585

QY 209 -ERDILAAALSHPLVTGLDQETRTKTLILILELCSSELDRLYRKGV-VTEAEVKYIQ 266
DB 586 QEISIMNCLHHPKLVQCVDAPFEKANIWMVLEIVSGGELFERIIDEDFELTERECIKYR 645

QY 267 QLVGGLHYLHSHGVHLHDIKPSNIIWMVHPAREDIKICDFGPAQNITPAELQFSQYGSPEF 326
DB 646 QISEGEVIHKQGIYVHLDKPENINCVNKTGTRIKLIDFGLARLENAGSLKVLFGTPEF 705

QY 327 VSPEIIQNVPVSEASDIWAMGVISYLSLTCSPFAGESDRATLLNVLEGRVSNSSPMAAH 386
DB 706 VAPEVINVEPIGYATDMNSIGVICVILVSLSPFMGNDNETLANVTSATWDFDEAFDE 765

QY 387 LSEDAKDTKATLQAPQAPSAACLSHPWFLKSMPEAEAHFINTKQKELLARSRWQR 446
DB 766 ISDDAKDFILNLLKMDKNLDCTOCLQHPWLMKDTKMEAKKLSKDKMKYMARRWKQX 825

QY 447 SLMSYKSIILVMSIPEL-----LRGPPDSP 471
DB 826 TGNVRAIGRLSSMAMISGLGRKSGTGTSP 858

Search completed: April 23, 2004, 15:11:16
Job time : 99 secs